

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 04:42:22 ; Search time 8348 Seconds
(without alignments)
11183.624 Million cell updates/sec

Title: US-10-086-156-23
Perfect score: 2154
Sequence: 1 atgacgagggggttttgcg.....taataaaaaaaaaaaaaa 2154

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	1200.4	55.7	135044	9	AC006001	AC006001 Homo sapi
C 2	1160.6	53.9	194464	2	AC146119	AC146119 Pan trogl
C 3	1031.4	47.9	135044	9	AC006001	AC006001 Homo sapi
C 4	588.6	27.3	4807	6	BD183414	BD183414 Novel gen
C 5	587	27.3	1211	9	BC042482	BC042482 Homo sapi
C 6	587	27.3	2576	6	AX714361	AX714361 Sequence
C 7	587	27.3	2576	6	AK056631	AK056631 Homo sapi
C 8	565.2	26.2	1124	6	BD275557	BD275557 MOLECULES
C 9	433.6	20.1	3864	6	AB056802	AB056802 Macaca fa
C 10	387.8	18.0	680	6	AX575790	AX575790 Sequence
C 11	374	17.4	374	11	GS1111	GS1111 SHGC-80071
C 12	315.2	14.6	363	6	AX575786	AX575786 Sequence
C 13	293.6	13.6	188791	10	AC117185	AC117185 Mus muscu
C 14	293.6	13.6	257003	10	AC122830	AC122830 Mus muscu
C 15	268.8	12.5	1103	10	BC051544	BC051544 Mus muscu
C 16	257.4	11.9	4143	9	AK127790	AK127790 Homo sapi
C 17	217.6	10.1	667	11	BV029947	BV029947 S212P6093
C 18	214.8	10.0	246	6	AX887431	AX887431 Sequence
C 19	214.8	10.0	246	6	BD027041	BD027041 Sequence
C 20	177.4	8.2	194464	2	AC146119	AC146119 Pan trogl
C 21	172.4	8.0	321	6	AX575788	AX575788 Sequence
C 22	167.2	7.8	203685	10	AC122339	AC122339 Mus muscu
C 23	159.2	7.4	213729	2	AC116246	AC116246 Rattus no
C 24	159.2	7.4	216180	2	AC119707	AC119707 Rattus no
C 25	121.2	5.6	178538	2	AC134778	AC134778 Homo sapi
C 26	121.2	5.6	203192	9	AC103910	AC103910 Homo sapi
C 27	118.4	5.5	79776	2	AC139431	AC139431 Homo sapi
C 28	117.8	5.5	82659	2	AC135725	AC135725 Homo sapi
C 29	117.8	5.5	204940	2	AC069007	AC069007 Homo sapi
C 30	117.8	5.5	207611	9	AC091132	AC091132 Homo sapi
C 31	116.4	5.4	18534	9	HSSERCA1	Y15724 Homo sapien
C 32	116.4	5.4	111370	9	AC067815	AC067815 Homo sapi
C 33	116.4	5.4	134362	9	AC084853	AC084853 Homo sapi
C 34	116.4	5.4	158414	9	AC005940	AC005940 Homo sapi
C 35	116.2	5.4	179564	9	AC139677	AC139677 Homo sapi
C 36	116.2	5.4	198470	9	AC046170	AC046170 Homo sapi
C 37	116.2	5.4	200420	9	AC138645	AC138645 Homo sapi
C 38	115.8	5.4	73947	9	AC068152	AC068152 Homo sapi
C 39	115.8	5.4	88004	9	AC002316	AC002316 Homo sapi
C 40	115.8	5.4	140386	9	AC015921	AC015921 Homo sapi
C 41	115.8	5.4	164874	2	AC138223	AC138223 Homo sapi
C 42	115.2	5.3	183093	9	AC138688	AC138688 Homo sapi
C 43	115	5.3	28746	2	AC008647	AC008647 Homo sapi
C 44	115	5.3	101744	9	AL512783	AL512783 Human DNA
C 45	114.6	5.3	122168	9	AC127383	AC127383 Homo sapi

ALIGNMENTS

RESULT 1
AC006001/c
LOCUS AC006001 135044 bp DNA linear PRI 02-OCT-2003
DEFINITION Homo sapiens PAC clone RP4-756H11 from 7, complete sequence.
ACCESSION AC006001
VERSION AC006001.2 GI:5708496
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135044)
AUTHORS Sulston, J.E. and Wilson, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

99063792
9847074
2 (bases 1 to 135044)
Lamar, B., Le, T. and Wohldmann, P.
The sequence of Homo sapiens PAC clone RP4-756H11
Unpublished (2001)
3 (bases 1 to 135044)
Waterston, R. H.
Direct Submission
Submitted (22-NOV-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 135044)
Waterston, R. H.
Direct Submission
Submitted (07-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 135044)
Waterston, R. H.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 135044)
Waterston, R. H.
Direct Submission
Submitted (26-APR-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 135044)
Wilson, R.
Direct Submission
Submitted (02-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 8, 1999 this sequence version replaced gi:3907522.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: sapiens@watson.wustl.edu

Summary Statistics

Center project name: H_DJ0756H11

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/OTB/CHR7 , send
mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

SOURCE INFORMATION:
This clone was derived from human PAC library RPCI-4, prepared by
Pieter de Jong and coworkers at http://www.chori.org using the
method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc.
(http://www.genomesystems.com) or Research Genetics, Inc.

(http://www.resgen.com) ; or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of RP4-756H11
actual end is at base position 135044 of RP4-756H11.

FEATURES	source
Location/Qualifiers	1. .135044
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
	/clone="RP4-756H11"
	/clone_lib="RPCI-4"
repeat_region	59. .349
	/rpt_family="Alu"
repeat_region	369. .488
	/rpt_family="ERV1"
repeat_region	494. .800
	/rpt_family="Alu"
repeat_region	835. .1030
	/rpt_family="L2"
repeat_region	1357. .1436
	/rpt_family="Alu"
repeat_region	1523. .1828
	/rpt_family="Alu"
repeat_region	2038. .2314
	/rpt_family="Alu"
repeat_region	2328. .2631
	/rpt_family="Alu"
repeat_region	2686. .2781
	/rpt_family="Alu"
repeat_region	2785. .2927
	/rpt_family="L1"
repeat_region	2937. .2973
	/rpt_family="L1"
repeat_region	2974. .3228
	/rpt_family="L1"
repeat_region	3229. .3516
	/rpt_family="Alu"
repeat_region	3517. .3546
	/rpt_family="L1"
repeat_region	3547. .3793
	/rpt_family="MER2_type"
repeat_region	3794. .3823
	/rpt_family="L1"
repeat_region	3824. .4073
	/rpt_family="Alu"
repeat_region	4074. .4144
	/rpt_family="L1"
repeat_region	4156. .4643
	/rpt_family="L1"
repeat_region	4657. .4887
	/rpt_family="Alu"
repeat_region	4895. .5060
	/rpt_family="MIR"
repeat_region	5182. .5624
	/rpt_family="L1"
repeat_region	5625. .5919
	/rpt_family="Alu"
repeat_region	5920. .6087
	/rpt_family="L1"
repeat_region	6214. .6370
	/rpt_family="Alu"
repeat_region	6492. .6700
	/rpt_family="MIR"
repeat_region	7521. .7616
	/rpt_family="MIR"
repeat_region	7652. .7971
	/rpt_family="Alu"
repeat_region	8017. .8294
	/rpt_family="Alu"

```
repeat_region 8299..8437
/rpt_family="Alu"
repeat_region 8769..8930
/rpt_family="Alu"
repeat_region 8972..9065
/rpt_family="Alu"
repeat_region 9079..9126
/rpt_family="Alu"
repeat_region 9187..9229
/rpt_family="(TA)n"
repeat_region 9230..9250
/rpt_family="(TG)n"
repeat_region 9231..9549
/rpt_family="AT-rich"
repeat_region 9550..9572
/rpt_family="Alu"
repeat_region 9648..9678
/rpt_family="AT-rich"
repeat_region 9750..9791
/rpt_family="(A)n"
repeat_region 9808..10423
/rpt_family="Alu"
repeat_region 11152..11226
/rpt_family="L2"
repeat_region 11716..12011
/rpt_family="MIR"
repeat_region 12015..12333
/rpt_family="Alu"
repeat_region 12465..12583
/rpt_family="MIR"
repeat_region 13722..13853
/rpt_family="MER1_type"
repeat_region 14062..14287
/rpt_family="L2"
repeat_region 14461..14592
/rpt_family="MIR"
repeat_region 14669..14965
/rpt_family="Alu"
repeat_region 15509..15814
/rpt_family="Alu"
repeat_region 15827..16007

Query Match 55.7%; Score 1200.4; DB 9; Length 135044;
Best Local Similarity 99.8%; Pred. No. 1.8e-272;
Matches 1212; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 926 AAAGCGGCTGATTTGGAACTCCCATTTGTCAGAGAGACAGAACATGAGAGAAAT 985
DB 32238 AAAGCGGCTGATTTGGAACTCCCATTTGTCAGAGAGACAGAACATGAGAGAAAT 32179

QY 986 CCCCTGTCCAGTTGCTGAGAGATTTCCAACTTCCAGGCTAGAGGATTTCCATTGA 1045
DB 32178 CCCCTGTCCAGTTGCTGAGAGATTTCCAACTTCCAGGCTAGAGGATTTCCATTGA 32119

QY 1046 GATGGGTTTACGTTGATTTGAACTGTCAGCAGCTGTCCTGTTTGCATGGCAA 1105
DB 32118 GATGGGTTTACGTTGATTTGAACTGTCAGCAGCTGTCCTGTTTGCATGGCAA 32059

QY 1106 TTCTGACCTTTTATGGCAACACCCCTGGGACAAACCAGATTTGATGATGATCC 1165
DB 32058 TTCTGACCTTTTATGGCAACACCCCTGGGACAAACCAGATTTGATGATGATCC 31999

QY 1166 AAAGTAGAATTTCCAGACAGTCCAAACCAAGGTATCAAGTGATGTTTCCAGAGTGAAGG 1225
DB 31998 AAAGTAGAATTTCCAGACAGTCCAAACCAAGGTATCAAGTGATGTTTCCAGAGTGAAGG 31939

QY 1226 CTCTCAGCGTGTCCAGGATTTCTGGGTTTGTAAAGCAGTACTGCGCCATTTGTGACCCCTG 1285
DB 31938 CTCTCAGCGTGTCCAGGATTTCTGGGTTTGTAAAGCAGTACTGCGCCATTTGTGACCCCTG 31879

QY 1286 TTTTITACTAATCATTCCTGCTTTTATAGGACATGTTTATCCCATGCTCCCTGGCAAGGA 1345
DB 31878 TTTTITACTAATCATTCCTGCTTTTATAGGACATGTTTATCCCATGCTCCCTGGCAAGGA 31819
```

```
QY 1346 TCCAGATTTCAATAGCTGAAACCCCTGTATAGCTTTTCTCTATTCTGCTTACCCAA 1405
DB 31818 TCCAGATTTCAATAGCTGAAACCCCTGTATAGCTTTTCTCTATTCTGCTTACCCAA 31759

QY 1406 GACACACTTGAAACCCCTCAGTAAAGCTATAGAGAGGGCATAGCAGGGGAGCCTCTCC 1465
DB 31758 GACACACTTGAAACCCCTCAGTAAAGCTATAGAGAGGGCATAGCAGGGGAGCCTCTCC 31699

QY 1466 CTTGTTTCTACAGCTCCATGATGAGGGTTGACTGAGGCCAGCAATCCTTGTAGTGCA 1525
DB 31698 CTTGTTTCTACAGCTCCATGATGAGGGTTGACTGAGGCCAGCAATCCTTGTAGTGCA 31639

QY 1526 CAGTTGCAATATAATTAACAGTTTCAAGATCTAGAGGTACCTTTTGAAGAAACCCCTTCA 1585
DB 31638 CAGTTGCAATATAATTAACAGTTTCAAGATCTAGAGGTACCTTTTGAAGAAACCCCTTCA 31579

QY 1586 GGGATATCTATCCACAGTAGCTGAGCAGCAGCAAGTGAACCTGAGATTTTGACCCACAC 1645
DB 31578 GGGATATCTATCCACAGTAGCTGAGCAGCAGCAAGTGAACCTGAGATTTTGACCCACAC 31519

QY 1646 AATAAGGGGGGCCCATTCCTTTTCAAAATATTTTGGCTTTCAGAAATACACTTCAATACAT 1705
DB 31518 AATAAGGGGGGCCCATTCCTTTTCAAAATATTTTGGCTTTCAGAAATACACTTCAATACAT 31459

QY 1706 GCAAAATATTGAGAGATTAAACAGAAATTCAGCTCTTATGCTCTAACTGAGAGAGCCACTG 1765
DB 31458 GCAAAATATTGAGAGATTAAACAGAAATTCAGCTCTTATGCTCTAACTGAGAGAGCCACTG 31399

QY 1766 CAAGTTGCAAGTTAGTACCATGTCAGCAGAGCCAGCTGAATCCAGAGCTTCCCAA 1825
DB 31398 CAAGTTGCAAGTTAGTACCATGTCAGCAGAGCCAGCTGAATCCAGAGCTTCCCAA 31339

QY 1826 GTGGACACAGCGGGGACTATTCCTGATGTCCTCCACCAAGAGAGAAATGAGCTGAGGC 1885
DB 31338 GTGGACACAGCGGGGACTATTCCTGATGTCCTCCACCAAGAGAGAAATGAGCTGAGGC 31279

QY 1886 GCTCTTGTCTGCTCCCAATGCAATCCATGTCAGTGTGCAATTCAGCTGTGATGATGATGAT 1945
DB 31278 GCTCTTGTCTGCTCCCAATGCAATCCATGTCAGTGTGCAATTCAGCTGTGATGATGATGAT 31219

QY 1946 TGGCATTTTGGAACTTGTATCTGACATGTAAGACAGCCTTACATATGCGGTGGGTGTC 2005
DB 31218 TGGCATTTTGGAACTTGTATCTGACATGTAAGACAGCCTTACATATGCGGTGGGTGTC 31159

QY 2006 AGGGCTCACACTTGTATCTAGCAGCTTTTGGAGGCTGAGGTGGGAGATGCTTGAGC 2065
DB 31158 AGGGCTCACACTTGTATCTAGCAGCTTTTGGAGGCTGAGGTGGGAGATGCTTGAGC 31099

QY 2066 ACAGGAGTTCCAGACAGCCTGAGCAACATGGCGAAATCCTGTCTCTTCAAGAAATAA-A 2124
DB 31098 ACAGGAGTTCCAGACAGCCTGAGCAACATGGCGAAATCCTGTCTCTTCAAGAAATAA 31039

QY 2125 ATAATAATAATAAT 2138
DB 31038 ATAATAATAATAAT 31025

RESULT 2
AC146119
LOCUS
DEFINITION AC146119 194464 bp DNA linear HTG 01-AUG-2003
PROGRESS *** 31 unordered pieces.
ACCESSION AC146119
VERSION AC146119.1 GI:33387073
KEYWORDS HTG; HTGS PHASE1.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 194464)
AUTHORS Wilson,R.K.
TITLE The sequence of Pan troglodytes clone
```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 194464)
Wilson,R.K.
Direct Submission
Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Center project name: C_PT038003

----- Summary Statistics -----

Sequencing vector: M13, 0%

Sequencing vector: plasmid, 100%

Chemistry: Dye-primer ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 177293 bases at least Q40

Consensus quality: 179213 bases at least Q30

Consensus quality: 180602 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1329: contig of 1329 bp in length
1330
1429: gap of unknown length
1430
3053: contig of 1624 bp in length
3054
3153: gap of unknown length
3154
4975: contig of 1822 bp in length
4976
5075: gap of unknown length
5076
6209: contig of 1134 bp in length
6210
6309: gap of unknown length
6310
7326: contig of 1017 bp in length
7327
7426: gap of unknown length
7427
10326: contig of 2900 bp in length
10327
10426: gap of unknown length
10427
12412: contig of 1986 bp in length
12413
12512: gap of unknown length
12513
14266: contig of 1754 bp in length
14267
14366: gap of unknown length
14367
16448: contig of 2082 bp in length
16449
16548: gap of unknown length
16549
18757: contig of 2209 bp in length
18758
18857: gap of unknown length
18858
22270: contig of 3413 bp in length
22271
22370: gap of unknown length
22371
25054: contig of 2684 bp in length
25055
25154: gap of unknown length
25155
28295: contig of 3141 bp in length
28296
28395: gap of unknown length
28396
31482: contig of 3087 bp in length
31483
31582: gap of unknown length
31583
34542: contig of 2960 bp in length
34543
34642: gap of unknown length
34643
37879: contig of 3237 bp in length
37880
37979: gap of unknown length
37980
43297: contig of 5318 bp in length
43298
43397: gap of unknown length
43398
48231: contig of 4834 bp in length
48232
48331: gap of unknown length
48332
55882: contig of 7551 bp in length
55883
55982: gap of unknown length
55983
64790: contig of 8808 bp in length
64791
64890: gap of unknown length
64891
73831: contig of 8941 bp in length

* 73832 73911: gap of unknown length
* 73932 82314: contig of 8383 bp in length
* 82315 82414: gap of unknown length
* 82415 90706: contig of 8292 bp in length
* 90707 90806: gap of unknown length
* 90807 98797: contig of 7991 bp in length
* 98798 98897: gap of unknown length
* 98898 106599: contig of 7802 bp in length
* 106700 106799: gap of unknown length
* 106800 117151: contig of 10352 bp in length
* 117152 117251: gap of unknown length
* 117252 129326: contig of 12075 bp in length
* 129327 129426: gap of unknown length
* 129427 145201: contig of 15775 bp in length
* 145202 145301: gap of unknown length
* 145302 161680: contig of 16379 bp in length
* 161681 161780: gap of unknown length
* 161781 193229: contig of 31449 bp in length
* 193230 193329: gap of unknown length
* 193330 194464: contig of 1135 bp in length.

FEATURES

source

1. .194464

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/chromosome="UNK"

/clone="RP43-3803"

1. .1329

/note="assembly_name:Contig10"

1430. .3053

/note="assembly_name:Contig11"

3154. .4975

/note="assembly_name:Contig17"

5076. .6209

/note="assembly_name:Contig18"

6310. .7326

/note="assembly_name:Contig19"

7427. .10326

/note="assembly_name:Contig20"

10427. .12412

/note="assembly_name:Contig21"

12513. .14266

/note="assembly_name:Contig22"

14367. .16448

/note="assembly_name:Contig23"

16549. .18757

/note="assembly_name:Contig24"

18858. .22270

/note="assembly_name:Contig25"

22371. .25054

/note="assembly_name:Contig26"

25155. .28295

/note="assembly_name:Contig27"

28396. .31482

/note="assembly_name:Contig28"

31583. .34542

/note="assembly_name:Contig29"

34643. .37879

/note="assembly_name:Contig30"

37980. .43297

/note="assembly_name:Contig31"

43398. .48231

/note="assembly_name:Contig32"

48332. .55882

/note="assembly_name:Contig33"

55983. .64790

/note="assembly_name:Contig34"

64891. .73831

/note="assembly_name:Contig35"

73932. .82314

/note="assembly_name:Contig36"

82415. .90706

/note="assembly_name:Contig37"

```
misc_feature 90807..98797
/note="assembly_name:Contig38"
misc_feature 9898..106699
/note="assembly_name:Contig39"
misc_feature 106800..117151
/note="assembly_name:Contig40"
misc_feature 117252..129326
/note="assembly_name:Contig41"
misc_feature 129427..145201
/note="assembly_name:Contig42"
misc_feature 145302..161680
/note="assembly_name:Contig43"
misc_feature 161781..193229
/note="assembly_name:Contig44"
misc_feature 193330..194464
/note="assembly_name:Contig6"

ORIGIN
Query Match 53.9%; Score 1160.6; DB 2; Length 194464;
Best Local Similarity 98.7%; Pred. No. 4.8e-263;
Matches 1203; Conservative 0; Mismatches 9; Indels 7; Gaps 3;

QY 926 AAGGCGGCTGATTTGGAACTCCCATGTGCTCCAGAGAGACAGCAATGAGAGAAAT 985
DB 142260 AAGGCGGCTGATTTGGAACTCCCATGTGCTCCAGAGAGACAGCAATGAGAGAAAT 142319

QY 986 CCCCTGTCCAGTGGCTGCAGAGATATCCAACTTCATGGCTAGAGATTCCATTGA 1045
DB 142320 CCCCTGTCCAGTGGCTGCAGAGATATCCAACTTCATGGCTAGAGATTCCATTGA 142379

QY 1046 GATGGGTTTACGTTGATTTTGAACCTCTGACAGCTGTTCTCTGTTTCATGGCAA 1105
DB 142380 GATGGGTTTACGTTGATTTTGAACCTCTGACAGCTGTTCTCTGTTTCATGGCAA 142435

QY 1106 TTCTGACCCCTTTATGGCAACACACCCTGGGCAACCCAGATTGTAGATTGATGCC 1165
DB 142436 TTCTGACCCCTTTATGGCAACACACCCTGGGCAACCCAGATTGTAGATTGATGCC 142495

QY 1166 AAGGTAGAAATTCAGAGAGTCCAAACCAAGTATCAAGTATGTTTCCAGAGTGGAAAG 1225
DB 142496 AAGGTAGAAATTCAGAGAGTCCAAACCAAGTATCAAGTATGTTTCCAGAGTGGAAAG 142555

QY 1226 CTCTCACCGTGTCCAGGATTTCTGGGTTTGAAGCAGTACTGCCATTTGTGACCCCTG 1285
DB 142556 CTCTCACCGTGTCCAGGATTTCTGGGTTTGAAGCAGTACTGCCATTTGTGACCCCTG 142615

QY 1286 TTTTCTTACCTATATCTGCTTTTATAGGATGTTTATCCCATGCTCCCTGGGCAAGGA 1345
DB 142616 TTTTCTTACCTATATCTGCTTTTATAGGATGTTTATCCCATGCTCCCTGGGCAAGGA 142675

QY 1346 TCCAGAAATCCAAATAGCTGAAACCCCTGTTATAGCTTTTCTCTATTTCTGCTTACCCAA 1405
DB 142676 TCCAGAAATCCAAATAGCTGAAACCCCTGTTATAGCTTTTCTCTATTTCTGCTTACCCAA 142735

QY 1406 GACACATTTGAACCCCTCAGTAAAGCTATAGAGAGGCGCATGAGCAGGGGCGAGCTCTCC 1465
DB 142736 GACACATTTGAACCCCTCAGTAAAGCTATAGAGAGGCGCATGAGCAGGGGCGAGCTCTCC 142795

QY 1466 CTGTGTTCTAGAGCTCCATGATGAGGGTTGATGCTGAGGCGAGCAATCTTTGATGTTGA 1525
DB 142796 CTGTGTTCTAGAGCTCCATGATGAGGGTTGATGCTGAGGCGAGCAATCTTTGATGTTGA 142855

QY 1526 CAGTTGCAATATAATTAACAGTTTCAAGATCTAGAGGTACCTTTTGAAGAACCCCTTCA 1585
DB 142856 CAGTTGCAATATAATTAACAGTTTCAAGATCTAGAGGTACCTTTTGAAGAACCCCTTCA 142915

QY 1586 GGGATATCTATCCAGTAGCTGGAGCAGCAAGGTGAACCTGAGATTTTGAACCCACAC 1645
DB 142916 GGGATATCTATCCAGTAGCTGGAGCAGCAAGGTGAACCTGAGATTTTGAACCCACAC 142975

QY 1646 AATAAGGGGGGCCATTTCTTTTCAAAATATTTTGGCTTCAGAAATACATTTCAATACAT 1705
DB 142976 AATAAGGGGGGGCCATTTCTTTTCAAAATATTTTGGCTTCAGAAATACATTTCAATACAT 143035
```

```
QY 1706 GCAATATTTGAGAGATTAAAGAAATTCAGCTCTTATGCTCTTAACTGAGAAGGCACTG 1765
DB 143036 GCAATATTTGAGAGATTAACTCTTATGCTCTTAACTGAGAAGGCACTG 143093

QY 1766 CAAAGTTGAGTTAGTATGATCCATGTGAGAGAGGCGAGCTGAATCCAGAGTTCCTCAA 1825
DB 143094 CAAAGTTGAGTTAGTATGATCCATGTGAGAGAGGCGAGCTGAATCCAGAGTTCCTCAA 143153

QY 1826 GTGGACACAGCGGGGACTATTCCTGATGTCCACCCAGAGAGGAGAGTGTGAGGCG 1885
DB 143154 GTGGACACAGCGGGGACTATTCCTGATGTCCACCCAGAGAGGAGAGTGTGAGGCG 143213

QY 1886 GCTCTTGCTCTGCCCAAAATGCAATGTCATGTCATTCAGGTGTCACCCCATTTCAAAATAACA 1945
DB 143214 GCTCTTGCTCTGCCCAAAATGCAATGTCATTCAGGTGTCACCCCATTTCAAAATAACA 143273

QY 1946 TGGCAATTTTGAACCTTGTATCTGACATGTAAAGACCGCTTACATTTGGGGTGGGTGC 2005
DB 143274 TGGCAATTTTGAACCTTGTATCTGACATGTAAAGACCGCTTACATTTGGGGTGGGTGC 143333

QY 2006 AGGGCTCACACTTGTATCTTAGCATTTCGAGGCTGAGTGGGCGAGATTGCTTGAGC 2065
DB 143334 AGTGGCTCACACTTGTATCTTAGCATTTCGAGGCTGAGTGGGCGAGATTGCTTGAGC 143393

QY 2066 ACAGAGTTCCAGACCGCTGAGCAACATGCGGAAATCCTGTCTTCAAGAAATA-AA 2124
DB 143394 ACAGAGTTCCAGACCGCTGAGCAACATGCGGAAATCCTGTCTTCAAGAAATAACA 143453

QY 2125 ATAATAATAATAATAAA 2143
DB 143454 ATAATAATAATAATAAA 143472

RESULT 3
AC006001 135044 bp DNA linear PRI 02-OCT-2003
DEFINITION Homo sapiens PAC clone RP4-756H11 from 7, complete sequence.
ACCESSION AC006001
VERSION AC006001.2 GI:5708496
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135044)
AUTHORS Sulston,J.E. and Wilson,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 135044)
AUTHORS Lamar,B., Le,T. and Wohldmann,P.
TITLE The sequence of Homo sapiens PAC clone RP4-756H11
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 135044)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 135044)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 135044)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 135044)
```

Thu Oct 14 09:37:19 2004

us-10-086-156-23.rge

AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 7 (bases 1 to 135044)
 AUTHORS Wilson, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-OCT-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Aug 8, 1999 this sequence version replaced gi:3907522.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: sapiens@watson.wustl.edu

 Summary Statistics

 Center project name: H_DJ0756H11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/OTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION:
 This clone was derived from human PAC library RP4-756H11, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
 The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.
 VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP4-756H11
 actual end is at base position 135044 of RP4-756H11.

FEATURES

Source

1. .135044
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosomes="7"
 /map="7"
 /clone="RP4-756H11"
 /clone_lib="RP4-756H11"
 59. .349
 /rpt_family="Alu"
 369. .488
 /rpt_family="ERV1"
 494. .800
 /rpt_family="Alu"
 835. .1030
 /rpt_family="L2"
 1357. .1436

repeat_region
 /rpt_family="Alu"
 1523. .1828
 /rpt_family="Alu"
 2038. .2314
 /rpt_family="Alu"
 2328. .2631
 /rpt_family="Alu"
 2686. .2781
 /rpt_family="Alu"
 2785. .2927
 /rpt_family="L1"
 2937. .2973
 /rpt_family="L1"
 2974. .3228
 /rpt_family="L1"
 3229. .3516
 /rpt_family="Alu"
 3517. .3546
 /rpt_family="L1"
 3547. .3793
 /rpt_family="MER2_type"
 3794. .3823
 /rpt_family="L1"
 3824. .4073
 /rpt_family="Alu"
 4074. .4144
 /rpt_family="L1"
 4156. .4643
 /rpt_family="L1"
 4657. .4887
 /rpt_family="Alu"
 4895. .5060
 /rpt_family="MIR"
 5162. .5624
 /rpt_family="L1"
 5625. .5919
 /rpt_family="Alu"
 5920. .6087
 /rpt_family="L1"
 6214. .6370
 /rpt_family="Alu"
 6492. .6700
 /rpt_family="MIR"
 7521. .7616
 /rpt_family="MIR"
 7652. .7971
 /rpt_family="Alu"
 8017. .8294
 /rpt_family="Alu"
 8299. .8437
 /rpt_family="Alu"
 8769. .8930
 /rpt_family="Alu"
 8972. .9065
 /rpt_family="Alu"
 9079. .9126
 /rpt_family="(TA)n"
 9187. .9229
 /rpt_family="(TG)n"
 9230. .9250
 /rpt_family="AT-rich"
 9251. .9549
 /rpt_family="Alu"
 9550. .9572
 /rpt_family="AT-rich"
 9648. .9678
 /rpt_family="(A)n"
 9750. .9791
 /rpt_family="Alu"
 9808. .10423
 /rpt_family="L2"
 11152. .11226
 /rpt_family="MIR"

```
repeat_region 11716..12011
/rpt_family="Alu"
repeat_region 12015..12333
/rpt_family="Alu"
repeat_region 12465..12583
/rpt_family="MIR"
repeat_region 13722..13853
/rpt_family="MER1_type"
repeat_region 14062..14287
/rpt_family="L2"
repeat_region 14461..14592
/rpt_family="MIR"
repeat_region 14665..14965
/rpt_family="Alu"
repeat_region 15509..15814
/rpt_family="Alu"
repeat_region 15827..16007

Query Match 47.9%; Score 1031.4; DB 9; Length 135044;
Best Local Similarity 93.6%; Pred. No. 1.5e-232;
Matches 1143; Conservative 0; Mismatches 66; Indels 12; Gaps 6;

QY 926 AAAGCGCGTATTTGGAACTCCCATTTGGTCCAGAGAGACAGAACATGAGAGAAAT 985
DB 82276 AAAGGAGCTATATTTGGAACTCCCATTTGGTCCAGAGAGGCGAGAAATGAGAGAAAT 82335

QY 986 CCCCTGTCCAGTGTGCTGAGGAGTATCCAACTT-CATGGGCTAGAGGATTCATTG 1044
DB 82336 CCCCTGTCCAGTGTGCTGAGGAGTATCCAACTTCCATGGGCTAGAGGATTCATTG 82395

QY 1045 AGATGGGTTTACGTTTGAATTTGAACACCTGTGACGACCTGTTCTGTTTGCATGGCA 1104
DB 82396 AGATAGGGTTTACGACTTGTGTTGAACACCTGTGACCA-TATTCTGTTTGCATGGCA 82454

QY 1105 ATTCTGACCTTTTATGGCAACACACCCCTGGGACACCCAGATTTGTAGATTGAGATC 1164
DB 82455 ATTCTGACCTTTTATGGCAACAAATACCCCTGGGACACCCCTGGATTTGTGGACTGAGATC 82514

QY 1165 CAAAGGTAGAAATTTCCAGACAGTCAACCAAGGTATCAAGTATGTTTCCAGAGTGGAG 1224
DB 82515 CAAAGGTAGAAATTTCCAGACAGTCAACCA--TATCAAGTATGTTTCCAGAGTGGAG 82572

QY 1225 GCTCTCAGCGTCCAGAGTTTCTGGGTTTGTAAAGCAGTACTGGCCATTTGTGACCT 1284
DB 82573 GCTCTCAGAGTCTCCAGAGTTTCTGGGTTTATAAGCAGTACTGGCCATTTGTGACCT 82632

QY 1285 GTTTTATACCTAATCATCTGCTTTTATAGGACATGGTTTACCAGATCCCTGGCAAGG 1344
DB 82633 GTTTTATACCTAATCATCTGCTTTTATAGGATATGGGTTTACCAGATCCCTGGCAAGG 82692

QY 1345 ATCCAGAAATCCAAATAGCTGAAACCCCTGTTATAGCTTTCTCTTATTTCTGCTTACC 1404
DB 82693 ATCCAGCAATCCAAATAGCTGAAACCCCTG-TATATCTTTCTCTTATTTGCTTACC 82751

QY 1405 AGACACATTTGAACCCCTCAGTAAGCTTATAGAGAGGCGCATAGACAGGCGAGCTCTC 1464
DB 82752 AGACACATTTGAACCCCTCAGTAAGCTTATAGAGAGGCGCATAGACAGGCGAGCTCTC 82811

QY 1465 CTTGTTTCTACAGCTCCATGATGAGGGTTGACGTAGGCGCAGCAATCTTGTAGGTG 1524
DB 82812 CTTGTTTCTCAGCTCCATGATGAGGGTTGCGCTGAGGCGCAGCAATCTTGTAGGTG 82871

QY 1525 ACAGTTGCAATATAATTAACAGTTTCAAGATCTCAGAGTACTCTTTTGAAGAACCCCTTC 1584
DB 82872 ACAGTTGCAATATAATTAACAGTTTCAAGTTTCAAGGCTCAGGCAACCGTTTGAAGAACCCCTTC 82931

QY 1585 AGGGATATCTATCCACAGTAGCTTGAGCAGCAGCAAGGTGAACCTGAGATTTTGAACCCACA 1644
DB 82932 AGGGATGCTATCTACAGCAGCTTGAGCAGCAGCAAGGTGAACCTGAGATTTTGAACCCACA 82991

QY 1645 CAATAAGG---GGGGGCAATCTTTTCAATAATTTTGGCTTTCAGAAATACACTTCATTAC 1701
DB 82992 CAATAAGGGTGGGGGGCAATCTTTTCAATAATTTTGGCTTTCAGCAGACACTTCATTAC 83051
```

```
QY 1702 ACATGCAAAATATTGAGAGATTAAACAGAAATTCAGCTCTTATGCTTAATCAGAGAGCC 1761
DB 83052 ATATGCAAAATATTGAGAGATTAAACAGAAATTCAGCTCTTATGCTTAATCAGAGAGCC 83111

QY 1762 ACTGCAAGTTGAGTAGGTACCCCATGTCGAGCAGAGGCGAGCTGAATCCAGAGCTTCC 1821
DB 83112 ACTGCAAGTTGAGTAGGTACCCCATGTCGAGCAGAGGCGAGCTGAATCCAGAGCTTCC 83171

QY 1822 CAAAGTGACACACAGCGGGGACTATTCTGATGTCCTCCACCCCAAGAGAGAAATGAGCTG 1881
DB 83172 CAAAGTGACACACAGCGGGGACTATTCTGATGTCCTCCACCCCAAGAGAGAAATGAGCTG 83231

QY 1882 AGGCGCTCTTGTCTGCTGCCCCAAATGCATCCCATGTCATTCAGCTGTCCACCATTCATCA 1937
DB 83232 TGGTGTCTGCTGCTGCCCCAAATGCATCCCATGTCATTCAGCTGTCCACCATTCATCA 83291

QY 1938 AAATAACATGGCATTTCTTGGAACTTGTATCTGATGTCATGTAAGACAGGCTACACATTTGG 1997
DB 83292 AAATAACATGGAAATTTCTTGGAACTTGTATCTGATGTAAGACAGGCTTACACATTTGG 83351

QY 1998 GTGGTGAGGAGGCTCACACTTGTAACTCTAGCATTGTGGAGGCTGAGGTGGGAGATT 2057
DB 83352 CTGGTGAGGAGGCTCACACTTGTAACTCTAGCATTGTGGAGGCTGAGGTGGGAGATT 83411

QY 2058 GCTTGAGCAGAGGAGTTCAGACAGGCTGAGCAACATGGCGAAATCCTGTCTCTTCAAG 2117
DB 83412 GCTTGAGCAGAGGAGTTCAGACAGGCTGAGCAACATGGCGAAATCCTGTCTCTTCAAG 83471

QY 2118 AAATAAATAATAATAATAAT 2138
DB 83472 AAATAAATAATAATAATAAT 83492

RESULT 4
BD183414 4807 bp DNA linear PAT 17-JUN-2003
LOCUS Novel genes and proteins encoded by the genes.
DEFINITION BD183414
ACCESSION BD183414
VERSION BD183414.1 GI:31875614
KEYWORDS JP 2002345492-A/127.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4807)
AUTHORS Ohara,O., Nagase,T. and Nakajima,D.
TITLE Novel genes and proteins encoded by the genes
JOURNAL Patent: JP 2002345492-A 127 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002345492-A/127
PD 03-DEC-2002
PF 26-FEB-2002 JP 2002049009
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,
A61P25/14,
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
FT CDS
Location/Qualifiers (60)..(989).
FEATURES
source 1. 4807
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 27.3%; Score 588.6; DB 6; Length 4807;
Best Local Similarity 96.2%; Pred. No. 4e-128;
Matches 603; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 230 CCGGCCAGCCCGCAGCGCGCGCTCATGTCAGCGCTGCTCGGCGAGCCAGAGGATGG 289
```



```
REFERENCE
AUTHORS      2 (bases 1 to 2576)
TITLE        Isegai,T., Otsuki,T. and Sugiyama,T.
JOURNAL      Submitted (24-OCT-2001) Takao Isegai, Helix Research Institute,
              Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
              (E-mail:genomics@hri.co.jp, Tel.81-438-52-3975, Fax:81-438-52-3986)
              NEDO human cDNA sequencing project supported by Ministry of
              Economy, Trade and Industry of Japan; cDNA full insert sequencing;
              Research Association for Biotechnology (RAB); cDNA library
              construction: Helix Research Institute (HRI) (supported by Japan
              Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
              HRI, and Biotechnology Center, National Institute of Technology and
              Evaluation; clone selection for full insert sequencing: RAB and
              HRI.

FEATURES
  source      1. .2576
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="OCBBF1000118"
              /tissue_type="brain"
              /clone_lib="OCBBF1"
              /dev_stage="fetus"
              /note="cloning vector: pME18SFL3"
  CDS         108..977
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="BAB71236.1"
              /db_xref="GI:16552087"
              /translation="MNVVTRGPSRDRQDGMSSDAEDDFLBPATPTAQAGHALPL
              LPQEPFVPLNIGAHFTLRLCYEDTILMAFSGRHYIPTDSEGRYFDRLDGT
              HFGDVLNLSGDI.PPFRVAVYKEAQYVAIGLPLEQLENMOPKKEVKRQAFGLGM
              PFKDHLRIEIVELARLAVOKARPAKLKVCVKEPDIPTPECPLNLSRPFESSED
              GOLFEHCEVDVSGPWEAVADVLLHCLVTLDSLAOGLTVDHQCIGVCDKHLVNHYY
              CKRPIYBFKITW"
  ORIGIN
    Query Match      27.3%; Score 587; DB 9; Length 2576;
    Best Local Similarity 96.0%; Pred. No. 8.6e-128;
    Matches 602; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

  QY 230 CCGGCCACCGCCGACCGCGCGCTCATGCGAGCGCTGCTCGGCGAGCCAGAGGGATGG 289
      |||||
  DB 52 CCGGGCCGCGCTCCGCCCGCGAGCGCGCCACTCGCCAGAGCCAGAGGGATGG 111
      |||||
  QY 290 TGGTAGTCACGGGGCGGAGCAGACACCGCTGCTCAGGACGGTGCCATGTCAGACTCTG 349
      |||||
  DB 112 TGGTAGTCACGGGGCGGAGCAGACACCGCTGCTCAGGACGGTGCCATGTCAGACTCTG 171
      |||||
  QY 350 ACGCCGAAGACGACTTCTTGAGCGCGGCACCGCGACCGCCACGAGCGGGGCGACGCGC 409
      |||||
  DB 172 ACGCCGAAGACGACTTCTTGAGCGCGGCACCGCGACCGCCACGAGCGGGGCGACGCGC 231
      |||||
  QY 410 TGCCCCCTGCTGCCACAGGAGTTTCTGAGGTGTTTCCCTTAAACATCGAGGGGCTCACT 469
      |||||
  DB 232 TGCCCCCTGCTGCCACAGGAGTTTCTGAGGTGTTTCCCTTAAACATCGAGGGGCTCACT 291
      |||||
  QY 470 TCACATACGCGCTGTCACATCGCGCTGTACGAGACACCATGTTGTCACGCCATGTTCA 529
      |||||
  DB 292 TCACATACGCGCTGTCACATCGCGCTGTACGAGACACCATGTTGTCACGCCATGTTCA 351
      |||||
  QY 530 GTGGGGCGGCACATACATCCCGACAGACTCCGAGGGCGGGTACTTTCATCGACCGAGATGGCA 589
      |||||
  DB 352 GTGGGGCGGCACATACATCCCGACAGACTCCGAGGGCGGGTACTTTCATCGACCGAGATGGCA 411
      |||||
  QY 590 CACACTTTGGAGATGTGTAATTTCTGCGCTCAGGGGACCTCCCAACCGAGGGCGGTG 649
      |||||
  DB 412 CACACTTTGGAGATGTGTAATTTCTGCGCTCAGGGGACCTCCCAACCGAGGGCGGTG 471
      |||||
  QY 650 TTCGAGCTGTGTACAAAGAGGCCCACTAGTACCTCGGGCCCTCTGAGAGCAGGTGG 709
      |||||
  DB 472 TTCGAGCTGTGTACAAAGAGGCCCACTAGTACCTCGGGCCCTCTGAGAGCAGGTGG 531
      |||||
```

```
QY 710 AGAACATGAGCCACTGAAGGGCGAGAGAGGTGCGCCAGCGTTTCTGGGACTCATGCCCT 769
      |||||
  DB 532 AGAACATGAGCCACTGAAGGGCGAGAGAGGTGCGCCAGCGTTTCTGGGACTCATGCCCT 591
      |||||
  QY 770 ATTACAAAGACCACTTGGAGCGGATTTGGAGATCGCCCGGCTGCGTGGGTCCAGCGGA 829
      |||||
  DB 592 ATTACAAAGACCACTTGGAGCGGATTTGGAGATCGCCCGGCTGCGTGGGTCCAGCGGA 651
      |||||
  QY 830 AGGCCCGCTTTCAGAGCTCAAGAGCT 856
      |||||
  DB 652 AGGCCCGCTTTCAGAGCTCAAGAGCT 678
      |||||

RESULT 8
LOCUS      BD275557          1124 bp          DNA          linear          PAT 17-JUL-2003
DEFINITION MOLECULES OF THE IMMUNE SYSTEM.
ACCESSION  BD275557
VERSION    BD275557.1 GI:33085325
KEYWORDS   JP 2002540791-A/2.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1124)
            Lu,D.A.M., Azimzai,Y., Baughn,M.R., Tang,T.Y., Lal,P. and Yue,H.
            MOLECULES OF THE IMMUNE SYSTEM
            Patent: JP 2002540791-A 2 03-DEC-2002;
            INCYTE PHARMACEUTICALS INC,Henry YUE,Preeti LAL,Tom Y TANG, Mariah
            R BAUGHN,Yalda AZIMZAI,Dyung Aina M LU
            OS Homo sapiens
            PN JP 2002540791-A/2
            PD 03-DEC-2002
            PF 04-APR-2000 JP 2000609571
            PR 05-MAY-1999 US 60/132647,05-APR-1999 US 60/127852 PI
            dyung aina m lu,yalda azimzai,mariah r baughn,tom y tang, PI
            preeti lal,
            PI henry yue
            CC This description about <220> can't be interpreted CC <220>
            CC <221> misc feature
            CC <223> Inctye ID No.: 2751129CB1
            FH Key Location/Qualifiers
            source      1. .1124
                       /organism="Homo sapiens"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:9606"
            ORIGIN
              Query Match      26.2%; Score 565.2; DB 6; Length 1124;
              Best Local Similarity 99.3%; Pred. No. 1.1e-122;
              Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

            QY 275 GAGCCAGAGGATGGTGTAGTCACTGAGCGGGGAGGAGCAGCGCTGTCAGGACGGTG 334
                |||||
            DB 17 GAGCCAGAGGATGGTGTAGTCACTGAGCGGGGAGGAGCAGCGCTGTCAGGACGGTG 76
                |||||
            QY 335 CCATGTCCAGCTCTGAGCGCCGAGAGACGACTTCTTGGAGCGGGCCACGCGCCAGCGCCACGC 394
                |||||
            DB 77 CCATGTCCAGCTCTGAGCGCCGAGAGACGACTTCTTGGAGCGGGCCACGCGCCAGCGCCACGC 136
                |||||
            QY 395 AGCGGGGCGCACGGCTGCCCTGCTGTCACAGAGTTTCTGAGGTGTTCCCTTAAACA 454
                |||||
            DB 137 AGCGGGGCGCACGGCTGCCCTGCTGTCACAGAGTTTCTGAGGTGTTCCCTTAAACA 196
                |||||
            QY 455 TCGGAGGGGCTCACTTTCACATACAGCTGTCCACACTGCGGTGTCTAGAGAGACACCATGT 514
                |||||
            DB 197 TCGGAGGGGCTCACTTTCACATACAGCTGTCCACACTGCGGTGTCTAGAGAGACACCATGT 256
                |||||
            QY 515 TGGCAGGCACTGTTTCAGTGGGGCGGCATACATCCCGACAGACTCCGAGGGCGGTACTTCA 574
                |||||
            DB 257 TGGCAGGCACTGTTTCAGTGGGGCGGCATACATCCCGACAGACTCCGAGGGCGGTACTTCA 316
                |||||
```


Lexicon Genetics Incorporated (US)
Location/Qualifiers
1. .680
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 18.0%; Score 387.8; DB 6; Length 680;
Best Local Similarity 88.2%; Pred. No. 8.2e-81;
Matches 456; Conservative 1; Mismatches 3; Indels 57; Gaps 1;
Qy 146 CCGGTACAGGCCACAGCTGGCGGCGAGCGGTTCGGCTTGAGGAGCACCACCGCCCTCCCGC 205
Db 1 CCGGTACAGGCCACAGCTGGCGGCGAGCGGTTCGGCTTGAGGAGCACCACCGCCCTCCCGC 60
Qy 206 CTGGGCACATGCTCTCGCCCTCCCTCGCGCCAGCCGCGAGCGCGCGCTCATGCCAGGCG 265
Db 61 CTGGGCACATGCTCTCGCCCTCCCTCGCGCCAGCCGCGAGCGCGCGCTCATGCCAGGCG 120
Qy 266 CTGCTCGGC----- 274
Db 121 CTGCTCGGCTAGGAGTGCCCGGGCGCGCTCTCGCCCGCCGAGCGCGCCAC 180
Qy 275 -----GAGCCAGAGGATGGTGTAGTACAGGGCGGGAGCCAGACAGCCCTCCTCAGG 328
Db 181 TGCCCGACAGCCAGAGGATGGTGTAGTACAGGGCGGGAGCCAGACAGCCCTCCTCAGG 240
Qy 329 ACAGTGGCATGTCAGCTCTGACCCGAGACCACTTTCTGGAGCCGCGCCACCGCAGCG 388
Db 241 ACAGTGGCATGTCAGCTCTGACCCGAGACCACTTTCTGGAGCCGCGCCACCGCAGCG 300
Qy 389 CCAGCGAGGGGCGACAGCGCTGCGCCCTGCTGTCACAGAGTTCTCTGAGGTGTTCCCG 448
Db 301 CCAGCGAGGGGCGACAGCGCTGCGCCCTGCTGTCACAGAGTTCTCTGAGGTGTTCCCG 360
Qy 449 TTAACATCGAGGGGCTCACTTCACTACAGCCCTGTCACAGCTGCTGCTACGAGACA 508
Db 361 TTAACATCGAGGGGCTCACTTCACTACAGCCCTGTCACAGCTGCTGCTACGAGACA 420
Qy 509 CCATGTTGGCAGCCATGTTCAAGTGGGCGGCACTACATCCCCACAGACTCCGAGGCGCGT 568
Db 421 CCATGTTGGCAGCCATGTTCAAGTGGGCGGCACTACATCCCCACAGACTCCGAGGCGCGT 480
Qy 569 ACTTCAICGACCGAGATGGACACACTTTGAGATGT 605
Db 481 ACTTCAICGACCGAGATGGACACACTTTGAGATGT 517

RESULT 11
G51111/c
LOCUS SHGC-80071 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION SHGC-80071 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G51111
VERSION G51111.1 GI:522288
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 374)
AUTHORS Olivier, M. and Cox, D.R.
TITLE Unpublished, Olivier, M., Cox, D.R. (2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: TCGGGTAAACCATGTCTCTAAAC

Primer B: GAGTATTCCAACTTCATGGGC
STS size: 325
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Amplifaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

BAC ends sequenced at TIGR from the RPC111 BAC library. Designed and developed at the Stanford Human Genome Center.

FEATURES
source

Location/Qualifiers
1. .374
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Human"
20. .344
20. .42
primer_bind
primer_bind
complement (322. .344)

ORIGIN

Query Match 17.4%; Score 374; DB 11; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.4e-77;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 977 AGAAGAAATCCCTGTCCAGTTCCCTGCGAGGAGTATTCACACACTTCATGGGCTAGAGGA 1036
Db 374 AGAAGAAATCCCTGTCCAGTTCCCTGCGAGGAGTATTCACACACTTCATGGGCTAGAGGA 315
Qy 1037 TTCCATTGAGATGGGGTTTACGTTCTTGATTTTGAACACCTGTGACACCTGTTCTCTGTTT 1096
Db 314 TTCCATTGAGATGGGGTTTACGTTCTTGATTTTGAACACCTGTGACACCTGTTCTCTGTTT 255
Qy 1097 GCATGGCAATTCGACCCCTTTTATGGCAACACACCCCTGGGACACCCAGATTGTAGA 1156
Db 254 GCATGGCAATTCGACCCCTTTTATGGCAACACACCCCTGGGACACCCAGATTGTAGA 195
Qy 1157 TTGAGATCCAAAGGTAGATTTCCACAGTCCCAACCAAGGTATCAAGTATGTTTCCAG 1216
Db 194 TTGAGATCCAAAGGTAGATTTCCACAGTCCCAACCAAGGTATCAAGTATGTTTCCAG 135
Qy 1217 AGTGAAGGCTCTCACCGTGTCCAGGATTTCTGGGGTTTGTAAAGCAGTACTGGCCATT 1276
Db 134 AGTGAAGGCTCTCACCGTGTCCAGGATTTCTGGGGTTTGTAAAGCAGTACTGGCCATT 75
Qy 1277 GTGACCCCTGTTTTTACCTTAATCATTTCTGTTTTTTAGGACATGGTTTACCCGATCCCT 1336
Db 74 GTGACCCCTGTTTTTACCTTAATCATTTCTGTTTTTTAGGACATGGTTTACCCGATCCCT 15
Qy 1337 GGCAGAGGATCCAG 1350
Db 14 GGCAGAGGATCCAG 1

RESULT 12
AX575786
LOCUS AX575786
DEFINITION Sequence 4 from Patent WO02068626.

AX575786 363 bp DNA linear PAT 07-JAN-2003
Sequence 4 from Patent WO02068626.

```

ACCESSION AX575786
VERSION AX575786.1 GI:27552274
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
1
Fridde,C.J., Gexhardt,B., Hilbun,E. and Turner,C.A.
Novel human ion channel-related proteins and polynucleotides
encoding the same
Patent: WO 0206826-A 4 06-SEP-2002;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
1..363
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 14.6%; Score 315.2; DB 6; Length 363;
Best Local Similarity 99.1%; Pred. No. 1.1e-63;
Matches 317; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 286 ATGGTGGTAGTCACGGGGCGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGC 345
Db 1 ATGGTGGTAGTCACGGGGCGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGC 60
QY 346 TGTGACGGCGAGACGACTTCTGAGCGGGCCAGCGCCAGCGCCAGCGCGGGGGCAC 405
Db 61 TGTGACGGCGAGACGACTTCTGAGCGGGCCAGCGCCAGCGCGGGGGCAC 120
QY 406 GGGTGGCCCTGCTCCACAGAGTTCTGAGGTTGTTCCCTTAACATCGAGGGGCT 465
Db 121 GGGTGGCCCTGCTCCACAGAGTTCTGAGGTTGTTCCCTTAACATCGAGGGGCT 180
QY 466 CACTTCACTACAGCCTGTCCACACTGCGGTGCTACGAGACACCATGTGCGACGCCATG 525
Db 181 CACTTCACTACAGCCTGTCCACACTGCGGTGCTACGAGACACCATGTGCGACGCCATG 240
QY 526 TTCAGTGGGGCGCATACATCCCGACAGACTCCGAGCGGCGGTACTTTCATGACCGAGAT 585
Db 241 TTCAGTGGGGCGCATACATCCCGACAGACTCCGAGCGGCGGTACTTTCATGACCGAGAT 300
QY 586 GGCACACACTTTGGAGATGT 605
Db 301 GGCACACACTTTGGATGT 320

RESULT 13
AC117185/c
LOCUS AC117185 188791 bp DNA linear ROD 05-NOV-2003
DEFINITION Mus musculus BAC clone RP23-37M7 from 8, complete sequence.
ACCESSION AC117185
VERSION AC117185.2 GI:21536193
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL
1
Haglund,K. and Schatzkammer,K.
The sequence of Mus musculus BAC clone RP23-37M7
Unpublished (2001)
REFERENCE
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
3 (bases 1 to 188791)
REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (08-APR-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

```

REFERENCE
AUTHORS
TITLE
JOURNAL

4 (bases 1 to 188791)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 188791)
Wilson,R.
Direct Submission

REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (05-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 21, 2002 this sequence version replaced gi:20069729.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics

Center project name: M_BA0037M07

COMMENT

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:

The RP23-23 BAC Library has been constructed by Kazutoyo Oseawa
and Minako Tatenio in the laboratory of Pieter de Jong
(http://www.chori.org) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (http://www.resgen.com) or
Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is
overlapped by AC122830.

FEATURES
source

Location/Qualifiers
1..188791
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP23-37M7"
/clone_lib="RP23-23"
6..531
/rpt_family="L1"
744..1012
/rpt_family="MaLR"
1386..2293
/rpt_family="L1"
2343..2550
/rpt_family="B2"
3654..3800
/rpt_family="B4"
3835..3997
/rpt_family="Alu"
4486..4563
/rpt_family="ERV1"

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

repeat_region 4936.5021
/rpt_family="Alu"
repeat_region 4955.5024
/rpt_family="B4"
repeat_region 7704.7831
/rpt_family="Alu"
repeat_region 8033.8179
/rpt_family="Alu"
repeat_region 9503.9572
/rpt_family="ID"
repeat_region 10671.10781
/rpt_family="Alu"
repeat_region 10859.11196
/rpt_family="L1"
repeat_region 11437.11614
/rpt_family="L1"
repeat_region 11693.11935
/rpt_family="L1"
repeat_region 12027.12252
/rpt_family="MaLR"
repeat_region 12335.12445
/rpt_family="L1"
repeat_region 12556.12640
/rpt_family="B4"
repeat_region 13164.13471
/rpt_family="L1"
repeat_region 14659.15026
/rpt_family="MaLR"
repeat_region 15779.16125
/rpt_family="L1"
repeat_region 16158.16313
/rpt_family="MaLR"
repeat_region 16314.16541
/rpt_family="L1"
repeat_region 16587.17116
/rpt_family="L1"
repeat_region 17143.17256
/rpt_family="L1"
repeat_region 17352.17578
/rpt_family="L2"
repeat_region 18997.19181
/rpt_family="B2"
repeat_region 19932.20138
/rpt_family="ERV1"
repeat_region 20186.20395
/rpt_family="B2"
repeat_region 20418.20567
/rpt_family="Alu"
repeat_region 21808.21855
/rpt_family="L1"
repeat_region 21984.22302
/rpt_family="L1"
repeat_region 22999.23076
/rpt_family="B4"
repeat_region 23077.23150
/rpt_family="ID"
repeat_region 23217.24022
/rpt_family="RMR1B"
repeat_region 24086.24219
/rpt_family="Alu"
repeat_region 25328.25749
/rpt_family="MaLR"
repeat_region 25981.26285
/rpt_family="MaLR"
repeat_region 26445.26691
/rpt_family="L1"
repeat_region 26695.26951
/rpt_family="L1"
repeat_region 27000.27172
/rpt_family="B4"
repeat_region 27010.27203
/rpt_family="B2"
repeat_region 27578.27807

repeat_region /rpt_family="RMR15"
27865.28310
/rpt_family="L1"
repeat_region 28497.28549
/rpt_family="ERV1"
repeat_region 28554.29041
/rpt_family="ERVK"
repeat_region 29627.29903
/rpt_family="L1"
repeat_region 29924.30062
/rpt_family="B4"
repeat_region 30385.30438
/rpt_family="ID"
repeat_region 30522.30797
/rpt_family="MaLR"
repeat_region 30977.31115
/rpt_family="Alu"
repeat_region 31160.31373
/rpt_family="L1"
repeat_region 31454.31544
/rpt_family="L1"
repeat_region 32135.32285
/rpt_family="B4"
repeat_region 32888.32959
/rpt_family="Alu"
repeat_region 33633.33736
/rpt_family="Alu"
repeat_region 33974.34383
/rpt_family="ERVK"
repeat_region 35030.35109
/rpt_family="RMR19B"
repeat_region 35110.35394

Query Match 13.6%; Score 293.6; DB 10; Length 188791;
Best Local Similarity 83.4%; Pred. No. 3.4e-58;
Matches 357; Conservative 0; Mismatches 69; Indels 2; Gaps 2;

Qy 429 GTTTCCTGAGTTGTTCCCTTAAACATCGAGGGGCTCACTTCACCTACACGCGCTGTCCAC 488
Db 96206 GTTTCCTGAGTTGTTCCCTTAAACATCGAGGGGCTCACTTCACCTACACGCGCTGTCCAC 96147

Qy 489 ACTGGGGTGTACGAAGACACCATGTTGGAGCCATGTTTCAGTGGGGGGGACTACATCCC 548
Db 96146 TCTGCAGCACTATGAAGACA-CATGCTGGTGTGTATGTTTAGAGGGCGGCACTTACATCCC 96088

Qy 549 CACAGACTCGAGGGCGGTTACTTTCATCGACCGAGATGGCACACACTTTGGAGATGTGCT 608
Db 96087 TCAGACTCCAGGATGGTACTTTCATCAACAGAGACACACATTTGGAGATGGCT 96028

Qy 609 GAATTTCTGCGCTCAGGGGACCTCCACCAGGAGCGTGTTCGAGCTGTGTACAAAGA 668
Db 96027 GAATTTCTGCGCTCAGGGGACCTCCACCAGGAGCGTGTTCGAGCTGTGTACAAAGA 95968

Qy 669 GGCCCGTACTATGCCATCGGGCCCTCTCTGGAGAGCTGGAGAAATGCGAGCCTGAA 728
Db 95967 GGTCCAGTACTATGCCATCGGGCCCTCTCTGGAGAGCTGGAGAAATGCGAGCCTGAA 95908

Qy 729 GGGGAGAGGTGGCCCAAGCGTTCTTGGAGCTCATGCCCTATTACAAAGACCACTTGA 788
Db 95907 GGGTGAAGGTGGCCCAAGCGTTCTTGGAGCTCATGCCCTATTACAAAGACCACTTGA 95849

Qy 789 GCGGATTTGGAGATGCCCGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 848
Db 95848 GCGGATTTGGAGATGCCCGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 95789

Qy 849 CAAGAGCT 856
Db 95788 GAAAGTCT 95781

RESULT 14
AC122830/c AC122830 257003 bp DNA linear ROD 08-NOV-2003
LOCUS

DEFINITION Mus musculus BAC clone RP23-224I23 from 8, complete sequence.
ACCESSION AC122830
VERSION AC122830.4 GI:29469632
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 257003)
AUTHORS Scott, K., Harkins, R., Cotton, M., Spalding, L., Creason, K., Mangiapanello, L. and Delaney, K.
TITLE The sequence of Mus musculus BAC clone RP23-224I23
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 257003)
AUTHORS Wilson, R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 257003)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 257003)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 257003)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 257003)
AUTHORS Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Apr 2, 2003 this sequence version replaced gi:22539363.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BA0224I23

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC117185 and AC116323.

FEATURES	Location/Qualifiers	source
repeat_region	1..257003 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="8" /map="8" /clone="RP23-224I23" /clone_lib="RPCI-23" 1..812 /rpt_family="L1"	
repeat_region	923..1089 /rpt_family="B2" complement(1012..1087) /product="tRNA-Ser" /note="Likely pseudogene (HMM Sc=18.64 / Sec struct Sc=4.38)"	
trna	1344..1615 /rpt_family="MaLR" 1669..1779 /rpt_family="L1" 1874..1967 /rpt_family="L1" 3001..3109 /rpt_family="Alu" 3187..3385 /rpt_family="B2" 3396..3594 /rpt_family="B2" 3602..3724 /rpt_family="BC1_MM" 3727..3832 /rpt_family="Alu" 4558..5185 /rpt_family="L1" 5186..5331 /rpt_family="MaLR" 5430..6192 /rpt_family="L1" 6291..6790 /rpt_family="L1" 6806..7044 /rpt_family="L1" 7821..7933 /rpt_family="L2" 8266..8352 /rpt_family="L1" 8372..9045 /rpt_family="L2" 9835..9956 /rpt_family="B4" 9976..10076 /rpt_family="Alu" 10351..10966 /rpt_family="MaLR" 11410..11602 /rpt_family="B2" 11622..11816 /rpt_family="L1" 12435..11625 /rpt_family="L1" 12771..12879 /rpt_family="Alu" 12892..13609 /rpt_family="L1" 13610..13823 /rpt_family="L1" 13830..14073 /rpt_family="L1"	


```

repeat_region 14361..14478
/rpt_family="Alu"
repeat_region 14542..14741
/rpt_family="B2"
repeat_region 14813..15193
/rpt_family="MaLR"
repeat_region 15194..15362
/rpt_family="MaLR"
repeat_region 15425..15923
/rpt_family="MaLR"
repeat_region 16265..16302
/rpt_family="MaLR"
repeat_region 16384..16540
/rpt_family="B4"
repeat_region 16597..16704
/rpt_family="Alu"
repeat_region 16785..16992
/rpt_family="B2"
repeat_region 20606..21525
/rpt_family="L1"
repeat_region 21523..21800
/rpt_family="L1"
repeat_region 21801..22146
/rpt_family="MaLR"
repeat_region 22147..22167
/rpt_family="L1"
repeat_region 22359..23677
/rpt_family="L1"
repeat_region 23820..23966
/rpt_family="Alu"
repeat_region 23973..24696
/rpt_family="L1"
repeat_region 24635..25435
/rpt_family="L1"
repeat_region 25441..25940
/rpt_family="L1"
repeat_region 26315..27000
/rpt_family="L1"
repeat_region 27001..27373
/rpt_family="MaLR"
repeat_region 27374..27429
/rpt_family="L1"
repeat_region 27640..27841
/rpt_family="ERVK"
repeat_region 27880..28183
/rpt_family="L1"
repeat_region 28861..28977
/rpt_family="MaLR"
repeat_region 29050..29397
/rpt_family="MaLR"
repeat_region 29504..29537
/rpt_family="Alu"
repeat_region 29550..29685
/rpt_family="Alu"

Query Match
Best Local Similarity 13.6%; Score 293.6; DB 10; Length 257003;
Matches 357; Conservative 0; Mismatches 69; Indels 2; Gaps 2;

QY 429 GTTCTCAGGTGTTTCCCTTAACATCGGAGGGGCTCACTTCACTACACCGCTGTCCAC 488
DB 55806 GTTCTCAGTTCGTCTCTTAACATTTGGAGGGGCTCACITTAACCGTGTATGCTTAC 55747

QY 489 ACTGGGTGCTACGACACACATGTTGGCAGCCATGTTCACTGGCGGCACTACATCCC 548
DB 55746 TCTGACGACACTATGAAGACA-CATGCTGGCTGTCAATGTTAGAGGGCGGCATTACATCCC 55688

QY 549 CACAGACTCCGAGGGCGGCTACTTTCATCGACCGAGATGGCCACACACTTTTGAGATGTGCT 608
DB 55687 TACAGACTCCAGGATTGGTACTTCATCAACAGACAGACACACACTTTTGAGATGGCT 55628

QY 609 GAAATTCCTGGCTCAGGGAGCTCCACCCACGAGGCGGTGTTGAGCTGTGTACAAAGA 668
DB 55627 GAACTTCTGCAATTCAGGGGAGCTGCCACCCGGGAGCAAGTGGCGAGCTGTGCACAAAGA 55568

```

```

QY 669 GGCCAGTACTATGCCATCGGCGCCCTCTCTGGAGCAGCTGGAGAACATCGACCACTGAA 728
DB 55567 GGTCCAGTACTATGCCATTTGGACCCCTCTCTGGAGCAGCTGGAGAACATCGACCACTGAA 55508

QY 729 GGCGGAGAGGTGCGCCCAAGCGCTTTCTGGGACTCATGCCCTATTACAAAGACCACTTGA 788
DB 55507 GGGTGAAGAGGTGCGCCAGACCTTCTCTGGGCTCATGTCC-ATTACAGGAGCACTTGA 55449

QY 789 GCGGATTGTGAGATCGCCCGGCTGCTGGTCCAGCGGAGGCGGCTTTGCCAAGCT 848
DB 55448 GCGGATAGTGGCGATTGCCAGCTTTCATACGATACAGCGGAGGCGGCTTTGCCAAGCT 55389

QY 849 CAAGAGCT 856
DB 55388 GAAAGTCT 55381

RESULT 15
LOCUS BC051544
DEFINITION Mus musculus potassium channel tetramerisation domain containing 7, mRNA (CDNA clone IMAGE:5254261), partial cds.
ACCESSION BC051544
VERSION BC051544.1 GI:30186133
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1103)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, C.M., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.B., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, M.A., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL MEDLINE
PUBMED 12477932
REFERENCE 2 (bases 1 to 1103)
AUTHORS Strausberg, R.
DIRECT SUBMISSION
SUBMITTED (28-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcaps-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (IULNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Brin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 113 Row: 1 Column: 15.

FEATURES

source

Location/Qualifiers

1..1103
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:5254261"
/tissue_type="Lung, 2 pooled mouse tumors."
/clone_lib="NCI CGAP_Lu33"
/lab_host="DH10B"
/note="Vector: pT7T3-Pac"
<1..1103
/gene="Kctd7"
/note="synonym: 4932409E18"
/db_xref="LocusID:212919"
/db_xref="MGI:2442265"
<1..608
/gene="Kctd7"
/codon_start=3
/product="Kctd7 protein"
/protein_id="AAH51544.1"
/db_xref="GI:30186134"
/db_xref="LocusID:212919"
/db_xref="MGI:2442265"
/translation="HEAEGRYFIDRGTGFGDVLNFLRSGDLPPEHVRAVHKRAQYV
AIGPLLEQLENQPLKGRVQAFGLMPYKDLHLEIVEIARLRAVORKARPAKLKV
CVFKEEMPITPYCEPLLSLRFERSESDGLFEHCEVDVDFGPEAVADVLLHCL
VTDLSAQGLTVDHQICVGDKHLNVNHYKRPYEFKITWM"
6..152
/gene="Kctd7"
/note="K tetra; Region: K+ channel tetramerisation domain.
The N-terminal, cytoplasmic tetramerisation domain (T1) of
voltage-gated K+ channels encodes molecular determinants
for subfamily-specific assembly of alpha-subunits into
functional tetrameric channels. It is distantly related to
the BTB/POZ domain pfam00651"
/db_xref="CDD:pfam02214"

gene

CDS

misc_feature

ORIGIN

Query Match : 12.5%; Score 268.8; DB 10; Length 1103;
Best Local Similarity 92.8%; Pred. No. 1.2e-52;
Matches 282; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 553 GACTCGAGGCGCGTACTTCATCGACCGAGATGSCACACACTTTGGAGATGCTGAAT 612
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6 GAGGCCGAGGCGCGTACTTCATCGATCGAGATGSCACACACTTTGGAGATGCTGAAC 65
QY 613 TTCCTGCGCTCAGGGACCTCCACCAGGAGCGTGTTCGAGCTGTGTACAAAGAGGCC 672
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
66 TTCCTGCGTTCAGGGACCTGCCGCCAGGGAGCATGTGCGAGCTGTGCACAGGAGGCC 125
QY 673 CAGTACTATGCCATCGGGCCCTCTCTGGAGCAGCTGGAGAACATCGACCACTGAAGGCC 732
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
126 CAGTACTATGCCATCGGGCCCTCTCTGGAGCAGCTGGAGAACATCGACCACTGAAGGCC 185
QY 733 GAGAAGGTGCGCAACGCTTCTTGGGACTCATGCCCTATTACAAAGACCACTTGGAGCGG 792
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
186 GAGAAGGTGCGCAGGCTTCTTGGGACTCATGCCCTATTACAAAGACCACTTGGAGCGG 245
QY 793 ATTGTGGAGATCGCCGCGTGGCTCCAGCGGAGGCCCGCTTTGCCAAGCTCAAG 852
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
246 ATTGTGGAGATTCGCCGCTTCGCGAGTACAGCGGAGGCCCGCTTTGCCAAGCTCAAA 305
QY 853 AGCT 856
Db |||
306 GTCT 309

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 08:34:58 ; Search time 1029 Seconds
(without alignments)
10612.096 Million cell updates/sec

Title: US-10-086-156-23
Perfect score: 2154
Sequence: 1 atgacgatggcggttttgcg.....taataaaaaaaaaaaaaa 2154

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues
Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2154	100.0	2154	15	US-10-086-156-23
2	587	27.3	2576	16	US-10-094-749-1045
3	585.4	27.2	1068	16	US-10-120-988-408
4	465	21.6	519	13	US-10-296-115-18
5	417	19.4	1100	13	US-10-027-632-31358
6	417	19.4	1100	16	US-10-027-632-31358
7	387.8	18.0	680	14	US-10-024-579-8
8	315.2	14.6	363	14	US-10-024-579-4
9	174	8.1	473	9	US-09-864-761-1385
10	174	8.1	583	15	US-10-086-156-29
11	170.4	8.0	321	14	US-10-024-579-6
12	172.4	7.9	173	9	US-09-864-761-18144
13	116.4	5.4	18534	13	US-10-240-425-1575
14	116.4	5.4	60430	13	US-10-087-192-1732

c 15	116.2	5.4	646	13	US-10-027-632-110966	Sequence 110966,
c 16	116.2	5.4	646	13	US-10-027-632-110967	Sequence 110967,
c 17	116.2	5.4	646	13	US-10-027-632-110968	Sequence 110968,
c 18	116.2	5.4	646	16	US-10-027-632-110966	Sequence 110966,
c 19	116.2	5.4	646	16	US-10-027-632-110967	Sequence 110967,
c 20	116.2	5.4	646	16	US-10-027-632-110968	Sequence 110968,
c 21	111	5.2	535	13	US-10-027-632-281283	Sequence 281283,
c 22	111	5.2	535	16	US-10-027-632-281283	Sequence 281283,
c 23	111	5.2	43871	17	US-10-741-601-5628	Sequence 5628, Ap
c 24	110.8	5.1	804	15	US-10-198-846-4375	Sequence 4375, Ap
c 25	110.6	5.1	535	13	US-10-027-632-281282	Sequence 281282,
c 26	110.6	5.1	535	16	US-10-027-632-281282	Sequence 281282,
c 27	110.4	5.1	2094	13	US-10-027-632-99288	Sequence 99288, A
c 28	110.4	5.1	2094	13	US-10-027-632-99288	Sequence 99288, A
c 29	110.4	5.1	2094	16	US-10-027-632-99288	Sequence 99288, A
c 30	110.4	5.1	2094	16	US-10-027-632-99289	Sequence 99289, A
c 31	110	5.1	87869	17	US-10-717-597-194	Sequence 194, App
c 32	109.8	5.1	1595	9	US-09-822-849A-250	Sequence 250, App
c 33	109.8	5.1	2027	16	US-10-104-047-100	Sequence 100, App
c 34	109.4	5.1	227931	16	US-10-085-117-274	Sequence 274, App
c 35	109.4	5.1	325348	16	US-10-085-117-358	Sequence 358, App
c 36	109.2	5.1	28667	16	US-10-264-237-2863	Sequence 2863, Ap
c 37	109.2	5.1	28668	16	US-10-264-237-2864	Sequence 2864, Ap
c 38	109	5.1	80420	17	US-10-322-281-42	Sequence 42, Appl
c 39	106.8	5.0	168198	17	US-10-322-696-88	Sequence 88, Appl
c 40	106.2	4.9	1246	13	US-10-027-632-122821	Sequence 122821,
c 41	106.2	4.9	1246	13	US-10-027-632-122822	Sequence 122822,
c 42	106.2	4.9	1246	16	US-10-027-632-122821	Sequence 122821,
c 43	106.2	4.9	1246	16	US-10-027-632-122822	Sequence 122822,
c 44	106.2	4.9	3139	13	US-10-027-632-115236	Sequence 115236,
c 45	106.2	4.9	3139	13	US-10-027-632-115237	Sequence 115237,

ALIGNMENTS

RESULT 1
US-10-086-156-23
; Sequence 23, Application US/10086156
; Publication No. US20030054989A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING TWO NOVEL HUMAN POTASSIUM CHANNEL BETA-SU
; FILE REFERENCE: D0115NP
; CURRENT APPLICATION NUMBER: US/10/086,156
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/272,190
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/274,258
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 23
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
US-10-086-156-23

Query Match 100.0%; Score 2154; DB 15; Length 2154;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACGATGGCGGTTTTCGGAATAGAAAGGGGAAAGGACCACTCAGGCGCGCGCG 60
Db 1 ATGACGATGGCGGTTTTCGGAATAGAAAGGGGAAAGGACCACTCAGGCGCGCGCG 60
Qy 61 CTGGCGCTGCTGCTCTTCGACTGGGCGAGCTTCTGCAATCAGGGCGGAACAGCGCG 120
Db 61 CTGGCGCTGCTGCTCTTCGACTGGGCGAGCTTCTGCAATCAGGGCGGAACAGCGCG 120

121 GCGTCGCCAGTAGCGGAGCGGTGCGGTGAGCCCGAGTCGGCGCGAGCGGGTCGGC 180
 121 GCGTCGCCAGTAGCGGAGCGGTGCGGTGAGCCCGAGTCGGCGCGAGCGGGTCGGC 180
 181 GTTCAGGAGGACACCGCCCTCCCGCCCTGCGCACTGCGCTCTCGCCCGCCCTCGGCGCAGCCC 240
 181 GTTCAGGAGGACACCGCCCTCCCGCCCTGCGCACTGCGCTCTCGCCCGCCCTCGGCGCAGCCC 240
 241 GCAGCGCGCGCTCATGCGAGGCGCTGCTCGCGGAGCGAGAGGAGATGGTGGTAGTCAAG 300
 241 GCAGCGCGCGCTCATGCGAGGCGCTGCTCGCGGAGCGAGAGGAGATGGTGGTAGTCAAG 300
 301 GGGCGGAGCGAGACAGCGCGTCTGAGAGCGGTGCGAGCGGTGCGAGCGGTGCGAGCGGTGCGAG 360
 301 GGGCGGAGCGAGACAGCGCGTCTGAGAGCGGTGCGAGCGGTGCGAGCGGTGCGAGCGGTGCGAG 360
 361 GACTTTCTGGAGCGCGCCAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 361 GACTTTCTGGAGCGCGCCAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 421 CCAAGAGGTTCTGAGGTTGTTCCCTTAAACATCGAGAGGCGGTCTACTTCACTACAGCG 480
 421 CCAAGAGGTTCTGAGGTTGTTCCCTTAAACATCGAGAGGCGGTCTACTTCACTACAGCG 480
 481 CTGTCCACACTGCGGTGCTAGAGAGACACCATGTTGCGAGCGCATGTTCACTGCGCGCGAC 540
 481 CTGTCCACACTGCGGTGCTAGAGAGACACCATGTTGCGAGCGCATGTTCACTGCGCGCGAC 540
 541 TACATCCCAACAGACTCCGAGGCGCGGTACTTCACTGAGAGGCGGTCTACTTCACTACAGCG 600
 541 TACATCCCAACAGACTCCGAGGCGCGGTACTTCACTGAGAGGCGGTCTACTTCACTACAGCG 600
 601 GATGTCTGAATTTCTGCGCTCAGGAGACCTCCACCGAGGAGCGGTGTTGAGCTGTG 660
 601 GATGTCTGAATTTCTGCGCTCAGGAGACCTCCACCGAGGAGCGGTGTTGAGCTGTG 660
 661 TACAAGAGGCGCGAGTACTGATGCGATCGCGGCGCTCTGAGAGCGGTGTTGAGAGCATGCG 720
 661 TACAAGAGGCGCGAGTACTGATGCGATCGCGGCGCTCTGAGAGCGGTGTTGAGAGCATGCG 720
 721 CCACTGAAGCGCGAGAGGTGCGCAAGCGTTCTGGGACTCATGCCCTATTACAAGAG 780
 721 CCACTGAAGCGCGAGAGGTGCGCAAGCGTTCTGGGACTCATGCCCTATTACAAGAG 780
 781 CACTTGAAGCGGATTTGGAGATCGCGCGCTGCGGTGCGGTGCGAGGAGCGCGCGCTTT 840
 781 CACTTGAAGCGGATTTGGAGATCGCGCGCTGCGGTGCGGTGCGAGGAGCGCGCGCTTT 840
 841 GCCAAGCTCAAGAGCTTGACACCTTCTGCGCTAATGAGTGTCTCATCAAGATGCCCTT 900
 841 GCCAAGCTCAAGAGCTTGACACCTTCTGCGCTAATGAGTGTCTCATCAAGATGCCCTT 900
 901 GGAGTCAATCATGATTAACGCAAGAGGCGGTGTTATTTGGAAACTCCCATTTGGTCCA 960
 901 GGAGTCAATCATGATTAACGCAAGAGGCGGTGTTATTTGGAAACTCCCATTTGGTCCA 960
 961 GAGAGACAGAAATGAGAGAAATCCCTGTCCAGTTGCGCTGCGAGGAGTATTCACAC 1020
 961 GAGAGACAGAAATGAGAGAAATCCCTGTCCAGTTGCGCTGCGAGGAGTATTCACAC 1020
 1021 TTCAATGGCTAGAGATTCATTTGAGATGGGGTTTACGCTTTGATTTTGAACACCTGTCA 1080
 1021 TTCAATGGCTAGAGATTCATTTGAGATGGGGTTTACGCTTTGATTTTGAACACCTGTCA 1080
 1081 GCATGTTCTCTGTTGATGGGATTCGAGATGGGGTTTACGCTTTGATTTTGAACACCTGTCA 1140
 1081 GCATGTTCTCTGTTGATGGGATTCGAGATGGGGTTTACGCTTTGATTTTGAACACCTGTCA 1140
 1141 AACCAGATTTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1200
 1141 AACCAGATTTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1200

1201 CAAAGTATGTTTCCAGAGTGAAGGCTCTCACGTTGCTCCAGGATTTCTGGGTTGTAA 1260
 1201 CAAAGTATGTTTCCAGAGTGAAGGCTCTCACGTTGCTCCAGGATTTCTGGGTTGTAA 1260
 1261 GCAGTACTGGCCATTTGTGACCCCTGTTTAACTAATCATCTGCTCTTTTATAGGACATG 1320
 1261 GCAGTACTGGCCATTTGTGACCCCTGTTTAACTAATCATCTGCTCTTTTATAGGACATG 1320
 1321 GTTTTACCCGATCCCTGGGAAAGATCCAGAAATCCAAATAGCTGAAACCCCTGTATAGC 1380
 1321 GTTTTACCCGATCCCTGGGAAAGATCCAGAAATCCAAATAGCTGAAACCCCTGTATAGC 1380
 1381 TTTTCTCTATTCTGCTTACCCCAAGACACACTTGAACCCCTCAGTAAGGCTATAGAGAG 1440
 1381 TTTTCTCTATTCTGCTTACCCCAAGACACACTTGAACCCCTCAGTAAGGCTATAGAGAG 1440
 1441 GGCATGAGCAGGGGAGCCTCTCCCTTGTGTTTCTAGCTTCCATGATGAGGGTGTGACTG 1500
 1441 GGCATGAGCAGGGGAGCCTCTCCCTTGTGTTTCTAGCTTCCATGATGAGGGTGTGACTG 1500
 1501 AGGCCAGCAATCTTCTGAGTGTGACAGTTGCAATATAATTAACAGTTTCAAGATCTAGA 1560
 1501 AGGCCAGCAATCTTCTGAGTGTGACAGTTGCAATATAATTAACAGTTTCAAGATCTAGA 1560
 1561 GGTACCTTTTGAAGAAACCCCTTTCAGGGATATCTATCCACAGTAGCTGAGCAGGCAAG 1620
 1561 GGTACCTTTTGAAGAAACCCCTTTCAGGGATATCTATCCACAGTAGCTGAGCAGGCAAG 1620
 1621 GTGAACCTGAGATTTTGAACCCACACAAATAGGGGGGCGCATTTCTTTTCAAAATTTTGG 1680
 1621 GTGAACCTGAGATTTTGAACCCACACAAATAGGGGGGCGCATTTCTTTTCAAAATTTTGG 1680
 1681 CTTTCAGAAATACCTTCAATACACATGCAAAATATTGAGAGATTAAACAGAAATTCAGACTCT 1740
 1681 CTTTCAGAAATACCTTCAATACACATGCAAAATATTGAGAGATTAAACAGAAATTCAGACTCT 1740
 1741 TATGCTTAACTGAGAGAGCCACTGCAAGTTGCAAGTTAGTACCCATGTGTCAGCAGAGGC 1800
 1741 TATGCTTAACTGAGAGAGCCACTGCAAGTTGCAAGTTAGTACCCATGTGTCAGCAGAGGC 1800
 1801 CAGCTGATCCAGAGCTTCCCAAGTGGACACAGGGGGGACTATTCCTGATGTCGCCAC 1860
 1801 CAGCTGATCCAGAGCTTCCCAAGTGGACACAGGGGGGACTATTCCTGATGTCGCCAC 1860
 1861 CCAAGAGAGAAATGAGCTGAGCGCTCTTGTCTGTCGCCAAATGCATCCCATGTGCATT 1920
 1861 CCAAGAGAGAAATGAGCTGAGCGCTCTTGTCTGTCGCCAAATGCATCCCATGTGCATT 1920
 1921 CACGTGTACCCCATTTCAAAATAACATGGCAATTTTGGAACTTGTATCTGACATGTGAAGA 1980
 1921 CACGTGTACCCCATTTCAAAATAACATGGCAATTTTGGAACTTGTATCTGACATGTGAAGA 1980
 1981 CCAGCTTACATTTGGGTGGGTGCGGGGCTCACACTTCTTAATCTTAGCCTTTGGAAG 2040
 1981 CCAGCTTACATTTGGGTGGGTGCGGGGCTCACACTTCTTAATCTTAGCCTTTGGAAG 2040
 2041 GCTGAGTGGGAGATTTGCTTGGACACAGGAGTTCCAGACAGCTCTGAGCAACATGCGGA 2100
 2041 GCTGAGTGGGAGATTTGCTTGGACACAGGAGTTCCAGACAGCTCTGAGCAACATGCGGA 2100
 2101 AATCCTGTCTCTTCAAGAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATA 2154
 2101 AATCCTGTCTCTTCAAGAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATA 2154

RESULT 2
 US-10-094-749-1045
 ; Sequence 1045, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI

Db 524 AGAATGACGACCTGAAGGCGAGAGGTGCGCCAGGTTCTGGGACTATGCCCCCT 583
 Qy 770 ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCGGCTGCGTGGCGTCCAGCGGA 829
 Db 584 ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCGGCTGCGTGGCGTCCAGCGGA 643
 Qy 830 AGGCCCGCTTGGCAAGCTCAAGACT 856
 Db 644 AGGCCCGCTTGGCAAGCTCAAGACT 670

RESULT 4
 US-10-296-115-18
 ; Sequence 18, Application US/10296115
 ; Publication No. US20040053248A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq Inc
 ; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 784PCT
 ; CURRENT APPLICATION NUMBER: US/10/296,115
 ; PRIOR FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: US09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: US09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 1478
 ; SEQ ID NO 18
 ; LENGTH: 519
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-296-115-18

Query Match 21.6%; Score 465; DB 13; Length 519;
 Best Local Similarity 94.2%; Pred. No. 4.1e-130;
 Matches 483; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 224 CCCCTCCGGCCAGCCGCGCCGCGTCATGCCAGGCGTCTGGCGAGCCAGAG 283
 Db 7 CCCGGTGCAGATTCTCTCCGCCGCCCGAGAGCCGCCCACTGCCAGAGCCAGAG 66
 Qy 284 GGATGGTGTAGTCA CGGGGGGGAGCCAGACAGCGCTGTCAGGAGCGTSCCAATGCCA 343
 Db 67 GGATGGTGTAGTCA CGGGGGGGAGCCAGACAGCGCTGTCAGGAGCGTSCCAATGCCA 126
 Qy 344 GCTCTGACCGCAAGACACTTCTGGAGCGGCGCA CGCCAGCCAGCCAGCGCGGGC 403
 Db 127 GCTCTGACCGCAAGACACTTCTGGAGCGGCGCGCCAGCCAGCGCGCGGGC 186
 Qy 404 ACGGCTGCCCTGCTGCACAGGAGTTCTCTGAGTTGTTCCCTTAACATCGAGGGG 463
 Db 187 ACGGCTGCCCTGCTGCACAGGAGTTCTCTGAGTTGTTCCCTTAACATCGAGGGG 246
 Qy 464 CTCACCTCACTACAGCCCTGTCACACTGCGGTGCTACGAAGACACCAATGTCAGAGCCA 523
 Db 247 CTCACCTCACTACAGCCCTGTCACACTGCGGTGCTACGAAGACACCAATGTCAGAGCCA 306
 Qy 524 TGTTCACTGGGGGCACTACATCCCAAGACTCGAGGCGGCTACTTCAATCGAGCGAG 583
 Db 307 TGTTCACTGGGGGCACTACATCCCAAGACTCGAGGCGGCTACTTCAATCGAGCGAG 366
 Qy 584 ATGGCAGACACTTTGGAGATGTCTGAATTTCTCGCTGCTAGGGGACCTCCACCCAGGG 643
 Db 367 ATGGCAGACACTTTGGAGATGTCTGAATTTCTCGCTGCTAGGGGACCTCCACCCAGGG 426
 Qy 644 ACGGTGTTGAGCTGTGTACAAAGAGCCAGTACTATGCCATCGGGCCCTCTCTGGAGC 703
 Db 427 ACGGTGTTGAGCTGTGTACAAAGAGCCAGTACTATGCCATCGGGCCCTCTCTGGAGC 486
 Qy 704 ACGTGGAGACATGACGACCTGAAGGGCGAGA 736
 Db 487 ACGTGGAGACATGACGACCTGAAGGGCGAGA 519

RESULT 5
 US-10-027-632-31358
 ; Sequence 31358, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 31358
 ; LENGTH: 1100
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-31358

Query Match 19.4%; Score 417; DB 13; Length 1100;
 Best Local Similarity 100.0%; Pred. No. 2.5e-115;
 Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 926 AAAGCGGCTGTATTTGGAACTCCCATTTGGTCCAGAGACAGACAAATGAGAAGAAAT 985
 Db 684 AAAGCGGCTGTATTTGGAACTCCCATTTGGTCCAGAGACAGACAAATGAGAAGAAAT 743
 Qy 986 CCCTGTCCAGTTGCTTGCAGGAGTATTCACACATTCATGGGCTAGAGATTCATTGA 1045
 Db 744 CCCTGTCCAGTTGCTTGCAGGAGTATTCACACATTCATGGGCTAGAGATTCATTGA 803
 Qy 1046 GATGGGGTTTACCTCTGTATTTGAACCTGTCCAGCTGTCTCTGTTCATGCGCAA 1105
 Db 804 GATGGGGTTTACCTCTGTATTTGAACCTGTCCAGCTGTCTCTGTTCATGCGCAA 863
 Qy 1106 TTCTGACCCCTTTTATGGCAACAAACCCCTGGGCAACCCAGATTTGTAGATTGAGATCC 1165
 Db 864 TTCTGACCCCTTTTATGGCAACAAACCCCTGGGCAACCCAGATTTGTAGATTGAGATCC 923
 Qy 1166 AAAGGTAGAAATTTCCAGACAGTCCCAACCAAGGTATCAAGTATGTTCCAGAGTGGAGG 1225
 Db 924 AAAGGTAGAAATTTCCAGACAGTCCCAACCAAGGTATCAAGTATGTTCCAGAGTGGAGG 983
 Qy 1226 CTCTCACCGTGTCCAGAGATTTCTGGGGTTTGTAAAGCAGTACTGGCCATTTGTACCCCTG 1285
 Db 984 CTCTCACCGTGTCCAGAGATTTCTGGGGTTTGTAAAGCAGTACTGGCCATTTGTACCCCTG 1043
 Qy 1286 TTTTACCTAACTCTCTCTTTTGTAGACATGTTTACCCGATTCCTCGGCAAA 1342
 Db 1044 TTTTACCTAACTCTCTCTTTTGTAGACATGTTTACCCGATTCCTCGGCAAA 1100

RESULT 6
 US-10-027-632-31358
 ; Sequence 31358, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.


```
;
;
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31358
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-31358

Query Match          19.4%; Score 417; DB 16; Length 1100;
Best Local Similarity 100.0%; Pred. No. 2.5e-115;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 926 AAAGCGCGTGTATTGGAAACTCCATTGGTCCAGAGAGACAGAAATGAGAAAT 985
Db 684 AAAGCGCGTGTATTGGAAACTCCATTGGTCCAGAGAGACAGAAATGAGAAAT 743

QY 986 CCCCTGTCCAGTTCCTGAGAGATATCCAACTTATGCGGTAGAGATTCATTGA 1045
Db 744 CCCCTGTCCAGTTCCTGAGAGATATCCAACTTATGCGGTAGAGATTCATTGA 803

QY 1046 GATGGGTTTACGCTTGTGATTTGAACACTGTGACGACTGTCTCTGTTTCATGGCAA 1105
Db 804 GATGGGTTTACGCTTGTGATTTGAACACTGTGACGACTGTCTCTGTTTCATGGCAA 863

QY 1106 TTCTGACCTTTTATGGCAACACCCCTGGGCAACCCAGATTTGTAGATTGAGATCC 1165
Db 864 TTCTGACCTTTTATGGCAACACCCCTGGGCAACCCAGATTTGTAGATTGAGATCC 923

QY 1166 AAAGGTAGAAATTTCCAGACAGTCCAAACCAAGGTATCAAGTATGTTTCCAGAGTGAAGG 1225
Db 924 AAAGGTAGAAATTTCCAGACAGTCCAAACCAAGGTATCAAGTATGTTTCCAGAGTGAAGG 983

QY 1226 CTCTCACCGTGTCCAGGATTTCTGGGTTTCTAAGCAGTACTGCGCATTTGTGACCCCTG 1285
Db 984 CTCTCACCGTGTCCAGGATTTCTGGGTTTCTAAGCAGTACTGCGCATTTGTGACCCCTG 1043

QY 1286 TTTTACCTAATCAATCTCTGCTTTTATGACATGTTTATCCCGATCCCTGGCAAA 1342
Db 1044 TTTTACCTAATCAATCTCTGCTTTTATGACATGTTTATCCCGATCCCTGGCAAA 1100

RESULT 7
US-10-024-579-8
; Sequence 8, Application US/10024579
; Publication No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 363
; TYPE: DNA
```

```
;
;
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 680
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-024-579-8

Query Match          18.0%; Score 387.8; DB 14; Length 680;
Best Local Similarity 88.2%; Pred. No. 1.4e-106;
Matches 456; Conservative 1; Mismatches 3; Indels 57; Gaps 1;

QY 146 CGGGTCAGSCCCAGCTGGCGCGAGCGGGTGGGGTTCGAGGAGCCACCGCCCTCCCGC 205
Db 1 CGGGTCAGSCCCAGCTGGCGCGAGCGGGTGGGGTTCGAGGAGCCACCGCCCTCCCGC 60

QY 206 CTGCGCACTGCTCTCGCCCCCTCTCGCGCCAGCCCGCCAGCCGCGCTCATGCCAGGCG 265
Db 61 CTGCGCACTGCTCTCGCCCCCTCTCGCGCCAGCCCGCCAGCCGCGCTCATGCCAGGCG 120

QY 266 CTGCTCGGC----- 274
Db 121 CTGCTCGCGGTAGGGAGTGCCTCGCGCGCGCGYCTCCGCCCGCCGAAAGCCGCGCCAC 180

QY 275 -----GAGCCAGAGGATGGTGTAGTACGCGGGGGGAGGACAGACAGCCCTCGTCAGG 328
Db 181 TGCCAGAGCCAGAGAGGATGGTGTAGTACGCGGGGGGAGGACAGACAGCCCTCGTCAGG 240

QY 329 ACGGTGCCATGTCCAGCTCTGACGCGCGAAGACGACTTTCTGGAGCGGGCCACGCGACGG 388
Db 241 ACGGTGCCATGTCCAGCTCTGACGCGCGAAGACGACTTTCTGGAGCGGGCCACGCGACGG 300

QY 389 CCACSCAGCGGGGCGACCGGCTGCGCTGCTGCTGCCACAGAGTTTCTGAGGTTGTTCCCG 448
Db 301 CCACSCAGCGGGGCGACCGGCTGCGCTGCTGCTGCCACAGAGTTTCTGAGGTTGTTCCCG 360

QY 449 TTAACATCGGAGGGGCTCACTACACGCTGTCCACACTGCGGTGCTACGAGACA 508
Db 361 TTAACATCGGAGGGGCTCACTACACGCTGTCCACACTGCGGTGCTACGAGACA 420

QY 509 CCATGTTGGCAGCCATGTTTCAGTGGCGCGCACTACATCCCCACAGACTCCGAGGGCGGT 568
Db 421 CCATGTTGGCAGCCATGTTTCAGTGGCGCGCACTACATCCCCACAGACTCCGAGGGCGGT 480

QY 569 ACTTCATCGACCGAGATGCGCACACTTTTGGAGATGT 605
Db 481 ACTTCATCGACCGAGATGCGCACACTTTTGGATGT 517

RESULT 8
US-10-024-579-4
; Sequence 4, Application US/10024579
; Publication No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 363
; TYPE: DNA
```


Query Match 8.1%; Score 174; DB 15; Length 583;
Best Local Similarity 83.2%; Pred. No. 1.1e-41;
Matches 199; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 368 TGGAGCGCGCCACGCGCGCCACGCGCGGGGCGCGCGTGCCTGCTGCCACAGG 427
DB 145 TGGAGCGCGCCACGCGTCTCATTCCTGCTGCTGAGAGCGCCCTGCTGTTCTTCC 204
QY 428 AGTTTCCTGAGGTGTTCCCTTAAACATCGAGGGGCTCACTTCACTACACGCTGTCCA 487
DB 205 AGTTTCCTGAGGTGTTCCCTTAAACATCGAGGGGCTCACTTCACTACACGCTGTCCA 264
QY 488 CACTGCGGTGCTACGAAGACACCATGTTGGAGCCCATGTTCACTGAGGGGCGGCACTACATCC 547
DB 265 CACTGCGGTGCTACGAAGACACCATGTTGGAGCCCATGTTCACTGAGGGGCGGCACTACATCC 324
QY 548 CCACAGACTCCGAGGGCGGCTACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
DB 325 CCACGAGCTCCGAGGGCGGCTACTTCATCGACCGAGATGGCACACACTTTGGGATGT 382

RESULT 11
US-10-024-579-6
; Sequence 6, Application US/10024579
; Publication No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins
; FILE REFERENCE: Lex-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 321
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-024-579-6

Query Match 8.0%; Score 172.4; DB 14; Length 321;
Best Local Similarity 82.8%; Pred. No. 2.3e-41;
Matches 197; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 368 TGGAGCGCGCCACGCGCGCCACGCGCGGGGCGCGCGTGCCTGCTGCCACAGG 427
DB 41 TGGAGCGCGCCACGCTCTCATTCCTGCTGCTGAGAGCCCTGCTGTTCTTCC 100
QY 428 AGTTTCCTGAGGTGTTCCCTTAAACATCGAGGGGCTCACTTCACTACACGCTGTCCA 487
DB 101 AGTTTCCTGAGGTGTTCCCTTAAACATCGAGGGGCTCACTTCACTACACGCTGTCCA 160
QY 488 CACTGCGGTGCTACGAAGACACCATGTTGGAGCCCATGTTCACTGAGGGGCGGCACTACATCC 547
DB 161 CACTGCGGTGCTACGAAGACACCATGTTGGAGCCCATGTTCACTGAGGGGCGGCACTACATCC 220
QY 548 CCACAGACTCCGAGGGCGGCTACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
DB 221 CCACGAGCTCCGAGGGCGGCTACTTCATCGACCGAGATGGCACACACTTTGGGATGT 278

RESULT 12
US-09-864-761-18144/c
; Sequence 18144, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18144
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006001.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.4
; OTHER INFORMATION: EST HUMAN HIT: A1674184.1, EVALUATE 5.00e-93
; OTHER INFORMATION: NT HIT: AJ001455.1, EVALUATE 1.60e+00
; OTHER INFORMATION: SWISSPROT HIT: Q13829, EVALUATE 1.00e-06
US-09-864-761-18144

Query Match 7.9%; Score 170.4; DB 9; Length 173;
Best Local Similarity 99.4%; Pred. No. 6.3e-41;
Matches 171; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 428 AGTTTCCTGAGGTGTTCCCTTAAACATCGAGGGGCTCACTTCACTACACGCTGTCCA 487
DB 172 AGTTTCCTGAGGTGTTCCCTTAAACATCGAGGGGCTCACTTCACTACACGCTGTCCA 113

QY 488 CACTGGGTCTACGAAGACACCATGTTGGCAGCCATGTTTCAGTGGGGCGGCACTACATCC 547
 Db 112 CACTGGGTCTACGAAGACACCATGTTGGCAGCCATGTTTCAGTGGGGCGGCACTACATCC 53
 QY 548 CCACAGACTCCGAGGCGCGTACTTTCATCGACCGAGATGGCACACACTTTGG 599
 Db 52 CCACGAGTCCGAGGCGCGTACTTTCATCGACCGAGATGGCACACACTTTGG 1

RESULT 13
 US-10-240-425-1575
 ; Sequence 1575, Application US/10240425
 ; Publication No. US20040033502A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Amanda
 ; APPLICANT: Boland, Joseph F.
 ; APPLICANT: Lord, Reginald V.
 ; APPLICANT: Alvarez, Chris
 ; APPLICANT: Wetzel, Jon C.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Vockley, Joseph G.
 ; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
 ; FILE REFERENCE: 44921-5026
 ; CURRENT APPLICATION NUMBER: US/10/240,425
 ; CURRENT FILING DATE: 2002-09-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/09847
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: US 60/193,446
 ; PRIOR FILING DATE: 2000-03-31
 ; NUMBER OF SEQ ID NOS: 1588
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1575
 ; LENGTH: 18534
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20040033502A1 Y15724
 US-10-240-425-1575

Query Match 5.4%; Score 116.4; DB 13; Length 18534;
 Best Local Similarity 75.8%; Pred. No. 3.2e-23;
 Matches 144; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 1965 TATCTGACATGTAAAGACCGCTACACATTCGGGTGGGTGCGAGGGCTCAGCTTGTAAAT 2024
 Db 2198 TATCAGACTTGAGATATAAATGTGAGCTGAGCGGGATGAGTGGCTTATACCTGTAAT 2247
 QY 2025 CCTAGCACTTTGGAAGGCTGAGGTGGGAGATTCCTTGAGCACAGGAGTTCAGACCCAGC 2084
 Db 2248 CCCAACATTTGGGAGGCGGAAGTGGGTGCGATCGCTTGAGTCCAGGAGTTTCAGACCCAGC 2307
 QY 2085 CTGAGCAATGGCGAAATCCTGCTCTCAAGAAATAAATAATAATAATAATAATAA 2144
 Db 2308 CTGAGCAATTTGGGAACCCCTGCTCTCAAAAAAATAAATAATAATAATAATAA 2367
 QY 2145 AAAAAAATAA 2154
 Db 2368 AAAAAAATAA 2377

RESULT 14
 US-10-087-192-1732
 ; Sequence 1732, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1732
 ; LENGTH: 60430
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-087-192-1732

Query Match 5.4%; Score 116.4; DB 13; Length 60430;
 Best Local Similarity 75.8%; Pred. No. 6.7e-23;
 Matches 144; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 1965 TATCTGACATGTAAAGACCGCTACACATTCGGGTGGGTGCGAGGGCTCAGCTTGTAAAT 2024
 Db 7741 TATCAGACTTGAGATATAAATGTGAGCTGAGCGGGATGAGTGGCTTATACCTGTAAT 7800
 QY 2025 CCTAGCACTTTGGAAGGCTGAGGTGGGAGATTCCTTGAGCACAGGAGTTCAGACCCAGC 2084
 Db 7801 CCCAACATTTGGGAGGCGGAAGTGGGTGCGATCGCTTGAGTCCAGGAGTTTCAGACCCAGC 7860
 QY 2085 CTGAGCAATGGCGAAATCCTGCTCTCAAGAAATAAATAATAATAATAATAATAA 2144
 Db 7861 CTGAGCAATTTGGGAACCCCTGCTCTCAAAAAAATAAATAATAATAATAATAA 7920
 QY 2145 AAAAAAATAA 2154
 Db 7921 AAAAAAATAA 7930

RESULT 15
 US-10-027-632-110966/c
 ; Sequence 110966, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 110966
 ; LENGTH: 646
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-110966

Query Match 5.4%; Score 116.2; DB 13; Length 646;
 Best Local Similarity 71.0%; Pred. No. 4.3e-24;
 Matches 154; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 QY 1931 CCATTCAAAATAACATGCGCATTCCTTGGAACCTTGATCTGACATGTGAAGACCCAGCTTACA 1990
 Db 276 CCTGTAATACTAAGTTTGATCACTTCAGTAAGGTGATCTTCCAGGTCTCTCCATTTGTAAT 217

This Page Blank (uspto)